Working Paper Series

Severance Pay in an Optimal Contract

WP 17-11R

Borys Grochulski
Federal Reserve Bank of Richmond

Tsz-Nga Wong
Federal Reserve Bank of Richmond

Yuzhe Zhang
Texas A&M University

FEDERAL RESERVE BANK OF RICHMOND®
Richmond ▪ Baltimore ▪ Charlotte

Severance pay in an optimal contract*

Borys Grochulski[1], Tsz-Nga Wong[1], and Yuzhe Zhang[2]

[1]Federal Reserve Bank of Richmond
[2]Texas A&M University

September 12, 2022

Abstract

We study the incentive role of severance compensation. In the canonical principal-agent model of Sannikov (2008), we introduce exogenous job destruction risk and show that compensation following job destruction can reduce the costs of incentives prior to job destruction. In an optimal contract, the award of severance suppresses the growth of the agent's value when this value is high, which eliminates the risk of inefficient retirement of the agent. To achieve this, however, severance compensation must exceed the agent's value conditional on job survival, effectively rewarding the agent for "bad luck" in the event of job destruction. High severance awards, thus, are a part of an optimal compensation package. Comparative statics with respect to job destruction risk offer a novel explanation for the positive wage-tenure profile observed in the data: average tenure and compensation should both be higher in jobs less exposed to job destruction risk.

Keywords: severance, incentive costs, job destruction, dynamic moral hazard principal-agent contracts, jump risk, wage-tenure profile

JEL codes: D86, J31

*The authors are grateful to Saki Bigio, David Fuller, Giovanni Gallipoli, Andreas Hornstein, Marianna Kudlyak, Jennifer La'O, B. Ravikumar, and Cheng Wang for their helpful comments. An earlier version of this paper was circulated under the title "Optimal incentive contracts with job destruction risk". The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System. Emails: borys.grochulski@rich.frb.org, russell.wong@rich.frb.org, zhangeager@tamu.edu.

1 Introduction

Standard theories of optimal incentives derive rich predictions on the level and performance-sensitivity of direct compensation but have much less to say about optimal use of severance compensation (Edmans and Gabaix, 2016). Severance pay is widely observed: 92 percent of businesses surveyed in 2018 reported offering some severance pay to their employees (Lee Hecht Harrison, 2018). In this paper, we add the risk of exogenous separation to a standard model of dynamic incentives (Sannikov, 2008) and study how severance pay, i.e., compensation conditional on separation, should be used in an optimal contract.

We find that awarding an agent compensation in the event of exogenous separation can reduce the cost of providing incentives to the agent prior to separation. Mechanically, by shifting the agent's compensation to the event of separation, severance pay decreases the value owed to him conditional on no separation. Therefore, severance pay will reduce the overall cost of providing incentives to the agent if this cost is increasing in the value owed to the agent prior to separation. In our model, as in Sannikov (2008), incentive costs are high when the agent's value prior to separation is either low or high; they are low when the agent's value is moderate. Therefore, it is desirable to reduce the value owed to the agent prior to separation only when this value is already high. An optimal contract, thus, awards severance compensation to agents who are owed a lot (e.g., those who occupy high-rank positions, typically in advanced stages of their careers) but does not provide any severance pay to agents who are owed little (e.g., those in low-rank positions, typically in early stages of their careers).

In our model, as in Sannikov (2008), the costs of incentives are high when the value owed to the agent—his so-called continuation value—is high because the agent's marginal utility of consumption is decreasing. Specifically, if the agent's continuation value reaches a particular threshold, the agent's marginal utility of consumption becomes low relative to his marginal disutility of effort. At that point, an optimal contract stops asking for effort altogether and the agent is retired, which is inefficient ex ante. In our model, severance pay mitigates this risk of inefficient retirement of the agent.

The event of exogenous separation, or job destruction, presents an opportunity to pay the agent without negative implications for the agent's incentives, because the relationship is at that point over. We model job destruction as an exogenous Poisson shock that remains outside of the agent's control. This event can represent, for example, an exogenous productivity shock or a large negative demand shock in the final product market that renders the relationship permanently unproductive.[1]

To mitigate the risk of inefficient retirement of the agent, severance pay must be sufficiently high. Indeed, how the risk of job destruction affects the agent's incentives prior to job

[1]Following Pissarides (1985) and Mortensen and Pissarides (1994), such shocks have been used extensively in the macro-labor literature.

destruction depends on whether the agent stands to lose or gain value at arrival of the Poisson job destruction shock. To suppress the growth of the agent's value prior to job destruction, which mitigates the risk of inefficient retirement, the agent must stand to gain at job destruction. For this reason, it is optimal to reward the agent at job destruction—with a severance award exceeding his value conditional on job survival—if the agent's value conditional on job survival is close enough to the inefficient retirement threshold.

Overall, an optimal contract can be divided into three phases. We refer to these phases as probation, early career, and late career. In late career, the agent's continuation value is high and the risk of inefficient retirement is paramount. Severance compensation is used to mitigate this risk, which means the value of the agent's severance award exceeds his value conditional on job survival, i.e., the agent gains value at arrival of a job destruction shock.

In the probation phase, the agent's value is close to zero. At zero, the agent does not respond to incentives at all, and thus the productive relationship must end, i.e., the agent is fired. To mitigate this risk, it would be beneficial to impose a negative severance, but the agent's limited liability makes this impossible. In the probation stage, thus, optimal severance pay is zero, i.e., the agent loses all continuation value if job destruction occurs in this phase of the contract.

The early career stage covers the middle region between probation and late career. In early career, the agent's continuation value is intermediate, and both endogenous separation risks—i.e., retirement and firing—are moderate. In this stage, the value of severance awarded to the agent is positive but remains below his continuation value conditional on job survival. At the point of transition between early and late career, the agent is indifferent to the event of job destruction: his severance exactly replaces his continuation value conditional on job survival.

As in Sannikov (2008), the profit of the principal/firm in an optimal contract is a hump-shaped function of the agent's value. Its peak is located at the cusp between the probation and early career stages of an optimal contract, which is where it is optimal for the firm to initiate the contract.

Job destruction is a bad-news shock in our model: the firm always loses value at job destruction. In an optimal contract, the firm's loss is loaded on late-career job destruction events. Indeed, in late career the risk of inefficient retirement of the agent is most severe. The firm mitigates this risk by awarding high severance compensation to the agent, effectively rewarding the agent for "bad luck." If job destruction arrives in this stage of the contract, thus, the firm incurs a large loss. This loss can be thought of as a payment for the extension of the relationship's duration achieved by the slowing down of the growth of the agent's continuation value near the inefficient retirement threshold. More generally, it can be thought of as an ex post payment for reducing the ex ante costs of incentives.

It is important for our results that the severance-triggering job destruction shock be completely exogenous, i.e., not indicative of any actions taken by the agent.[2] The COVID-19 pandemic, along with the lockdowns introduced to contain it, provides a clear example of such an exogenous negative shock, particularly for the high-contact industries like retail, entertainment, travel, and hospitality (Epiq, 2022). Consistent with our theory, several well-publicized corporate bankruptcy filings in the United States in mid-2020, including JCPenny and Hertz, were immediately preceded by large payouts to executives (Spencer, 2021). These filings stand in contrast to many examples of bankruptcies brought about by poor performance rather than a purely exogenous event, in which executives are typically dismissed without severance, e.g., the Lehman Brothers bankruptcy in 2008.

We should note that severance in our model is not merely a benefit awarded to highly-compensated agents simply because the value owed to them is high.[3] Rather, severance is a tool for reducing overall incentive costs that is effective when the incentive costs are increasing in the agent's continuation value. If incentive costs never increase with the agent's value, severance is not useful at all. This point is well illustrated by the analysis of Hoffmann and Pfeil (2010), where the implications of an exogenous "bad luck" shock are examined in a model in which there is no risk of inefficient retirement of the agent: the optimal severance following a job-destruction shock causing the firm's liquidation is always zero. The same result is obtained in Anderson et al. (2018): with incentive costs decreasing in the agent's value, it is never optimal to award severance to a departing agent.

In contrast to several other studies of the impact of exogenous jump risk on incentives, e.g., Piskorski and Tchistyi (2010) and Li (2017), we focus on job destruction risk, i.e., there is no further incentive problem to solve after the jump shock in our model. This assumption gives to the continuation value owed to the agent right after the jump shock the meaning of severance, which we study in this paper. This assumption also allows us to contrast the risk of endogenous versus exogenous separation. Clearly, the possibility of exogenous separation induced by a job destruction shock reduces the ex ante probability of endogenous separation at both the agent-firing threshold and the retirement threshold. In an optimal contract with severance, however, the risk of inefficient retirement is reduced much more comprehensively than the risk of inefficient firing of the agent.

Specifically, we show that, in probation and in the early-career stage, an optimal contract gives an upward drift to the agent's continuation value process. In these two stages, due to positive volatility of the contract, the agent faces the risk of hitting the firing threshold. Upon reaching the late-career stage, however, the contract becomes close to being stationary (has near-zero drift) with reduced volatility, which effectively "slows down" its dynamics in this region. The contract is thus unlikely to leave the late-career stage before

[2] DeMarzo et al. (2014) and Wong (2019) study models with negative jump risk that is partially controlled by the agent.

[3] This would indeed be true in a model with full information (first best), decreasing marginal utility of consumption, and exogenous separation shocks. In such a model, however, severance would be strictly positive at all levels of the continuation value.

arrival of a job destruction shock. In particular, the transition to the agent retirement threshold becomes very unlikely. This way, the risk of job destruction completely replaces the risk of inefficient agent retirement. We show this result analytically in the limit as the rate of arrival of job destruction becomes high, and also verify it numerically in a parametrized example with a reasonably low rate of arrival of job destruction.

Comparative statics with respect to the intensity of the Poisson shock allow us to quantify the impact of job destruction risk on the firm's ex ante profit and on the degree of compensation backloading achieved by an optimal contract. Higher risk of job destruction limits the possibility of backloading, which decreases the firm's ex ante profit.

Consistent with this comparative statics result, our analysis predicts that higher average job durations, higher wages, and higher severance should be observed in sectors, occupations, or localities that are less exposed to the risk of job destruction. Our model, thus, suggests that a positive wage-tenure profile observed in the data, as well as the severance-tenure profile, can be explained purely by the lower overall incentive costs achieved in more stable long-run relationships. This novel explanation for a positive wage-tenure profile is independent of the standard explanation related to accumulation of match-specific human capital, suggesting a different channel through which contract duration can affect observed compensation levels.

Severance has been studied in the literature on optimal contracts for CEOs. This literature finds a role for severance in settings in which the principal (the board of directors) faces additional frictions in addition to the CEO's moral hazard problem. In Almazan and Suarez (2003), the board has a time-consistency problem, and severance can be an ex ante useful commitment device dissuading the board from replacing the CEO with a marginally better one ex post. In Inderst and Mueller (2010), the CEO has private information about her outside options, and severance persuades the CEOs with weak outside options to quit. In He (2012), the CEO has access to private savings and can hurt the firm's liquidation value. Severance is a part of the CEO's compensation, even though the CEO is fired for poor performance, to preserve consumption smoothing. In our model, in contrast, moral hazard is a sole contracting friction, which makes our analysis also applicable to many jobs below the top executive positions. Severance mitigates the risk of inefficient retirement of the agent, which is not a focus of the literature on CEO incentives.

Organization The rest of this paper is organized as follows. Section 2 lays out an optimal contracting problem with job destruction risk, and Section 3 casts it in a recursive form. Section 4 examines first-order conditions. Section 5 characterizes the agent's exposure to job destruction risk in an optimal contract, while Section 6 does the same for the firm. Section 7 provides comparative statics with respect to the arrival rate of job destruction. Section 8 computes the expected contract duration and exit probabilities. Section 9 studies the limit when the arrival rate of the job destruction shock goes to infinity. Section 10 discusses testable implications of our model. Section 11 concludes.

2 A dynamic incentive problem with job destruction risk

We study the canonical dynamic principal-agent model of Sannikov (2008) extended to allow for exogenous job destruction and severance pay. Upon arrival of a job destruction shock, the productivity of the match jumps to zero permanently, i.e., the job is destroyed and it is optimal for the parties to separate. We model the exogenous job destruction shock as a Poisson time θ arriving with intensity $\lambda \geq 0$.

We also allow for endogenous separations. In particular, it is without loss of generality to restrict attention to a class of endogenous separations that take place when the agent becomes sufficiently poor or sufficiently rich. Formally, a separation of this kind takes place when the value owed to the agent reaches either 0 or a threshold W_{gp} defined in (13). Following Sannikov (2008, Section 3), we will refer to these events, respectively, as agent's firing and retirement.[4] The time of endogenous separation will be denoted by τ. By $\hat{\tau} \equiv \min\{\theta, \tau\}$ we denote the time of separation (exogenous or endogenous).

After separation, the agent can still be compensated by the firm, i.e., the agent can collect pension or severance pay. In the event of endogenous separation, the value owed to the agent is either 0 or W_{gp}, where W_{gp} is the agent's retirement value or pension. Severance represents the value promised to the agent in the event of exogenous job destruction. Specifically, we will refer to the continuation utility promised to the agent conditional on job destruction occurring at time t as the agent's severance value at t, and denote it by J_t.

The firm's continuation profit at separation from an agent who is owed value J will be denoted by $F_{\text{sep}}(J)$. This function is exogenous to the relationship between the firm and the agent; Section 2.1 provides further discussion.

Let X_t denote the cumulative output up to date t produced in the relationship. Before separation, i.e., for $t < \hat{\tau}$, as in the standard model, X_t follows

$$dX_t = A_t^a dt + \sigma dZ_t,$$

where $A_t^a \in \mathcal{A}$ is the agent's action (effort) at date t, and Z_t is a standard Brownian motion on $(\Omega, \mathcal{F}, P)$ independent of the job destruction shock θ. We assume that the set of feasible actions $\mathcal{A}$ is an interval $[0, \bar{A}]$ for some $\bar{A} > 0$.[5] After separation, i.e., for all $t \geq \hat{\tau}$, the process X_t follows

$$dX_t = 0,$$

i.e., no further output is produced in the match.

A contract $\mathcal{C}$ consists of a stopping time τ and three progressively measurable processes:

$$\mathcal{C} \equiv \left(\tau \geq 0, \{(A_t, C_t, J_t), 0 \leq t \leq \infty\}\right),$$

[4]The interpretation of endogenous separation at W_{gp} as agent retirement is consistent with models embedding dynamic incentive problems into an external labor market, e.g., Wang (2011).

[5]Our analysis can be extended to allow $\mathcal{A}$ to be any compact subset of $\mathbb{R}_+$.

where A_t is the action recommended for the agent to take at t, C_t is the agent's flow compensation, and J_t is the agent's promised severance value, i.e., his continuation value conditional on job destruction occurring at t. At each t, the agent chooses privately his action $A_t^a \in \mathcal{A}$ to maximize his utility. A contract is incentive compatible if $A_t^a = A_t$ at all t, i.e., the actual action chosen by the agent is that recommended by the contract.

The agent's expected payoff from an incentive-compatible contract is

$$\mathbb{E}\left[r\int_0^{\hat{\tau}} e^{-rt}\left(u(C_t)-h(A_t)\right)dt + e^{-r\hat{\tau}}\left(\mathbf{1}_{\{\hat{\tau}=\theta\}}J_{\hat{\tau}} + \mathbf{1}_{\{\hat{\tau}=\tau\}}\mathbf{1}_{\{W_\tau=W_{gp}\}}W_{gp}\right)\right],$$

where $r > 0$ and $\mathbf{1}_{\{.\}}$ is the indicator function. The agent's utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ is C^2 with $u' > 0$, $u'' < 0$, $\lim_{c\to 0} u'(c) = \infty$, and $u(0) = 0$. The agent's effort disutility function $h : \mathcal{A} \to \mathbb{R}_+$ is C^2 with $h' > 0$, $h'' > 0$, and $h(0) = 0$. In addition, we follow Sannikov (2008) in assuming

$$\lim_{a\to 0} h'(a) =: \gamma > 0.^{6} \tag{1}$$

Under an incentive compatible contract, at any $t < \hat{\tau}$, the agent's continuation value process is

$$W_t \equiv \mathbb{E}_t\left[r\int_t^{\hat{\tau}} e^{-r(s-t)}\left(u(C_s)-h(A_s)\right)ds + e^{-r\hat{\tau}}\left(\mathbf{1}_{\{\hat{\tau}=\theta\}}J_{\hat{\tau}} + \mathbf{1}_{\{\hat{\tau}=\tau\}}\mathbf{1}_{\{W_\tau=W_{gp}\}}W_{gp}\right)\right].$$

The agent's continuation value does not jump in an endogenous separation.[7] In an exogenous separation, however, it will be optimal to adjust the agent's continuation value discretely, i.e., $J_\theta \neq W_\theta$ *a.s.* We study the optimal jump in agent's continuation value at job destruction in detail in Section 5.

The firm's ex ante expected profit from an incentive compatible contract is

$$\mathbb{E}\Bigg[r\int_0^{\hat{\tau}} e^{-rt}(A_t - C_t)dt + e^{-r\hat{\tau}}\left(\mathbf{1}_{\{\hat{\tau}=\theta\}}F_{\text{sep}}(J_{\hat{\tau}}) + \mathbf{1}_{\{\hat{\tau}=\tau\}}\left(\mathbf{1}_{\{W_\tau=W_{gp}\}}F_{\text{sep}}(W_{gp}) + \mathbf{1}_{\{W_\tau=0\}}F_{\text{sep}}(0)\right)\right)\Bigg].$$

2.1 Severance cost and separation profit

In theory, the firm could keep the agent employed even after arrival of a job destruction shock θ. To focus on the interesting cases in which separation at job destruction is optimal,

[6] This assumption is convenient technically, as it implies that the volatility of the agent's continuation value remains bounded away from zero at all times prior to separation.

[7] To deliver promised value of 0 or W_{gp}, any adjustment to the agent's continuation value at endogenous separation would have to be mean-preserving spread. Such spreads are infeasible at 0 because of the agent's limited liability. They are suboptimal at W_{gp} under concave F_{sep}.

we will assume that the firm's separation payoff is at least as high as the payoff from keeping the agent employed after a job destruction shock has made the match permanently unproductive. That is, we will assume $F_{\text{sep}} \geq F_0$, where

$$F_0(W) \equiv -c \text{ such that } u(c) = W. \tag{2}$$

Here, F_0 represents the profit from keeping the agent employed after job destruction. Under F_0, the agent's continuation value W is delivered by constant compensation c with no effort required.[8] Note that F_0 is negative, strictly decreasing, and strictly concave.

The continuation profit function F_{sep} represents the firm's value from continuation after separation (e.g., the expected profit from searching for a replacement for the agent) less the cost to deliver to the agent his severance or retirement value at separation. Specifically, let

$$F_{\text{sep}}(J) \equiv D - rL \text{ such that } V(L) = J, \tag{3}$$

where $D \geq 0$ denotes the firm's value from continuation after separation, L is the lump-sum severance payment to the agent, and $V(L)$ denotes the agent's post-separation value function for any severance $L \geq 0$.

One example covered by (3) is the agent replacement model in Sannikov (2008, Section 3). There, $V(L) = u(rL)$ and $D = F(W_0) - \kappa$, where $F(W_0)$ is the profit from hiring a new agent with promised value W_0, and κ is a search cost. In this example, thus, $F_{\text{sep}}(J) = F(W_0) - \kappa + F_0(J)$. We will maintain flexible assumptions on F_{sep} to allow a wide class of models of the form (3).

Assumption 1 F'_{sep} *is strictly decreasing and weakly concave, with* $F'_{\text{sep}}(0) = 0$.

Assumption 2 $F'_{\text{sep}}(W) \geq F'_0(W)$ *for all* $W \leq W^*_{gp}$*, where* W^*_{gp} *is defined as the solution to* $F'_{\text{sep}}(W) = -1/\gamma$.

Assumption 2 ensures that the first derivative of the firm's profit function, $F'(W)$, is convex wherever $F'(W) \geq F'_{\text{sep}}(W)$. Together with the weak concavity of F'_{sep} in Assumption 1, this will ensure a single crossing between F' and F'_{sep} in an optimal contract.

3 Recursive formulation

In this section, we formulate the problem recursively and discuss the existence and computation of the solution.

[8] Indeed, permanent compensation $c = u^{-1}(W)$ and effort $a = 0$ deliver $W \geq 0$ to the agent because $h(0) = 0$ and $r \int_0^\infty e^{-rt} u(c) dt = u(c) = W$. The total cost to the firm to deliver W to the agent in this static way is $r \int_0^\infty e^{-rt} c dt = c$, hence the profit is $F_0(W) = -c \leq 0$.

3.1 Contract dynamics prior to job termination

Before separation, the dynamics of the agent's continuation value are

$$dW_t = r(W_t - u(C_t) + h(A_t))dt + rY_t(dX_t - A_t dt) + \Delta_t(dN_t - \lambda dt), \tag{4}$$

where $dX_t - A_t dt$ is the agent's observed performance relative to the benchmark $A_t dt$, rY_t represents the sensitivity of the agent's continuation value to his performance,

$$\Delta_t \equiv J_t - W_t \tag{5}$$

represents the sensitivity to the job destruction shock, and N_t is the counting process stopped at 1. That is, Δ_t is the jump in the agent's continuation value, i.e., the amount he gains, at arrival of a job destruction shock.

As in Sannikov (2008), a contract $\mathcal{C}$ is incentive compatible if and only if

$$A_t \in \operatorname*{argmax}_{a \in \mathcal{A}} Y_t a - h(a) \text{ at all } t < \hat{\tau}. \tag{6}$$

Note that interior effort $A_t \in (0, \bar{A})$ is incentive compatible if and only if

$$Y_t = h'(A_t). \tag{7}$$

At all $t < \theta$, we have $dN_t = 0$. Thus, under an incentive compatible contract, the dynamics of the agent's continuation value before separation, given in equation (4), are reduced to

$$dW_t = (rW_t - r(u(C_t) - h(A_t)) - \lambda \Delta_t)\, dt + r\sigma Y_t dZ_t. \tag{8}$$

The drift term in the above representation accounts for how W_t is delivered to the agent. The value owed to the agent grows at the rate of time preference less the flow of utility delivered to the agent, $r(u(C_t) - h(A_t))$, and less $\lambda\Delta_t$. The term $\lambda\Delta_t$ can be thought of a fair-odds premium charged to the agent in the event of no job destruction at t in order to account for the gain $\Delta_t = J_t - W_t$ the agent receives in the event of job destruction at t. The agent pays this premium indirectly through the drift of his continuation value W_t. Using this premium, and making the agent's severance value J_t history-dependent, a contract can better control the position of W_t while the relationship remains productive. Specifically, by setting J_t above W_t, the contract can charge a positive premium $\lambda\Delta_t$, which pushes W_t downward. Setting the severance value J_t below W_t has the opposite effect on the drift of W_t.

Since J_t does not enter the performance-sensitivity term, $r\sigma Y_t$, severance does not impact incentives directly but only indirectly by affecting the average position of W_t prior to separation. In this respect, severance compensation J_t is similar to immediate compensation c_t paid to the agent while the job remains active.

Example 1 Suppose the contract provides no severance, i.e., $J_t = 0$ or, equivalently, $\Delta_t = -W_t$, at all t. Since the agent loses all continuation value at job destruction, in order to deliver W_t to the agent prior to job destruction, the contract must provide a negative premium term $-\lambda W_t$, which pushes the agent's continuation value upward. This push is especially strong when W_t high, i.e., near the agent's retirement point W_{gp}, which accelerates the retirement of the agent and thus is inefficient.

Example 2 Suppose $J_t = W_t$ or, equivalently, $\Delta_t = 0$, at all t. With this severance promise, the agent is fully insured against job destruction risk. Since job destruction, as an exogenous event, carries no information about the agent's actions, not exposing the agent to this risk provides a reasonable benchmark. An optimal contract, however, will expose the agent to job destruction risk, i.e., $J_t \neq W_t$, at almost all t.

3.2 The contracting problem in recursive form

The firm designs a contract $\mathcal{C}$ to maximize its profit subject to the requirements of incentive compatibility and the agent's participation. Assuming the agent's reservation value of 0, the contract must deliver nonnegative value to the agent ex ante: $W_0 \geq 0$.

In the recursive form, the firm's problem is to maximize the profit value it can attain in the relationship with the agent, for any $W \geq 0$. Let us denote this value function by $F(W)$. The controls in this problem are c, a, Y, and J. With these controls, the drift and volatility of the agent's continuation value process, in (8), can be written as

$$\begin{aligned} \mu(W) &= rW - r(u(c) - h(a)) - \lambda(J - W), && (9) \\ \nu(W) &= r\sigma Y(W). && (10) \end{aligned}$$

The value function F must satisfy the following HJB equation

$$\begin{aligned} rF(W) = \max_{c,a,Y,J} \big\{ & ra - rc + F'(W)\left(rW - r(u(c) - h(a)) - \lambda(J - W)\right) + \frac{1}{2}F''(W)r^2\sigma^2Y^2 \\ & - \lambda(F(W) - F_{\text{sep}}(J))\big\}. \end{aligned} \tag{11}$$

This HJB equation is standard except for two terms that capture the job destruction shock. The term $-F'(W)\lambda(J-W)$ shows how much the firm values the fair-odds premium $\lambda(J-W)$ the agent is charged conditional on job survival. The term $-\lambda(F(W) - F_{\text{sep}}(J))$ captures the direct impact of job destruction on the firm's continuation profit: at job destruction, the firms gives up $F(W)$ and gains $F_{\text{sep}}(J)$.

3.3 Boundary conditions and existence of solution

The HJB equation (11) can be solved starting from $W = 0$ using the boundary condition

$$F(0) = F_{\text{sep}}(0) \tag{12}$$

and a conjectured initial slope $F'(0)$. In Appendix A.1, we show that the forward-shooting procedure of Sannikov (2008) pins down the initial slope $F'(0)$ and determines the optimal solution curve F, which represents the firm's value function in an optimal contract. In particular, barring degenerate cases excluded by Assumption 3 in Appendix A.1, the optimal solution curve F satisfies $F(W) \geq F_{\text{sep}}(W)$ for all $W \geq 0$, and there exists $W_{gp} > 0$ such that

$$F(W_{gp}) = F_{\text{sep}}(W_{gp}) \quad \text{and} \quad F'(W_{gp}) = F'_{\text{sep}}(W_{gp}). \tag{13}$$

The initial slope $F'(0)$ is strictly positive and the optimal solution curve is strictly concave with a unique maximum at $W^* \equiv \operatorname{argmax} F(W)$, where $0 < W^* < W_{gp}$. An optimal contract starts with $W_0 = W^*$ and follows (8), where A_t, Y_t, C_t, and J_t are determined by the policies that attain F in the HJB equation (11).[9] Since $F(W) = F_{\text{sep}}(W)$ at $W = 0$ and $W = W_{gp}$, endogenous separation is indeed optimal at either of these points.[10]

4 Optimal use of compensation

In this section, we use the first-order conditions with respect to J and c to discuss the optimal use of severance compensation and compare it to immediate compensation. Neither form of compensation affects the incentive constraint (7) directly. Instead, compensation pays down the value W_t owed to the agent and, thus, affects the costs of incentives indirectly via the average position of W_t.

The first-order condition with respect to J in the HJB equation (11) is

$$-F'(W)\lambda + \lambda F'_{\text{sep}}(J) \leq 0, \tag{14}$$

with equality if $J > 0$. By increasing the agent's severance value J, on the one hand, the firm increases its severance costs conditional on job destruction (i.e., reduces its separation profit) by $-F'_{\text{sep}}(J)$ on the margin. This cost is realized at the rate of arrival of job destruction, λ. On the other hand, by increasing the severance value J the firm increases the fair-odds premium charged to the agent conditional on the job's survival, which reduces the drift in the agent's value W_t, also at the rate λ. The first derivative of the value function, $F'(W)$, captures the impact of the drift of W_t on the firm's profit.

The first-order condition for c in the HJB equation (11) is

$$-F'(W)u'(c) \leq 1, \tag{15}$$

[9] Given the strictly positive marginal disutility of effort, (1), the agent's effort $A_t = a(W_t)$ and the volatility $Y_t = Y(W_t)$ of the agent's continuation value process remain bounded away from zero at all times prior to separation. Allowing for $\bar{A}$ sufficiently large, the agent's effort remains in $(0, \bar{A})$, i.e., volatility Y_t and effort A_t are related by the first-order condition (7), as long as $0 < W_t < W_{gp}$ and $t < \theta$.

[10] The endogenous separation time τ, thus, represents the first exit time of the agent's continuation value process W_t from the open interval $(0, W_{gp})$.

with equality if $c > 0$. A marginal unit of immediate compensation for the agent costs the firm 1 and decreases the drift of the agent's continuation value at the rate $u'(c)$.

If $W \in (0, W^*]$, then $F'(W) \geq 0$, i.e., a reduction in the drift of W_t actually lowers the firm's profit, as it increases the risk of reaching $W = 0$, where the agent must be fired. In this region, it is optimal to set the agent's compensation at zero: $c(W) = J(W) = 0$. Following Sannikov (2008), we will refer to this region of the state space as *probation*. In probation, to minimize the risk of inefficient termination at $W = 0$, the agent does not receive any direct compensation and is not entitled to any severance.

If $W \in (W^*, W_{gp})$, then $F'(W) < 0$, i.e., a reduction in the drift of W_t increases the firm's profit. In this region, optimal compensation is positive, in both forms, as determined by

$$F'_{\text{sep}}(J(W)) = \frac{-1}{u'(c(W))} = F'(W). \tag{16}$$

Proposition 1 *In an optimal contract, $c(W) = J(W) = 0$ for all $W \in (0, W^*]$. On (W^*, W_{gp}), $c(W)$ and $J(W)$ are strictly increasing, with $J(W) < W_{gp}$ for all $W \in (0, W_{gp})$.*

Proof In Appendix A.2. ■

In sum, at the low endogenous separation point $W = 0$, the agent is fired with no severance. At the high endogenous separation point $W = W_{gp}$, the agent is retired with value W_{gp}. In an exogenous separation at the moment of job destruction, $\theta < \tau$, we have $0 < J_\theta < W_{gp}$, with J_θ increasing in W_θ.

Figure 1 provides an illustration.[11] As noted in Example 2, J_θ is generally not equal to W_θ. In probation, $J(W) = 0$. Positive severance is awarded to the agent outside of probation. Next, we discuss the impact the agent's severance promise has on his continuation value prior to job destruction.

5 The agent's exposure to job destruction risk

In this section, we examine the agent's exposure to job destruction, which is captured by the jump in the agent's continuation value triggered by the arrival of a job destruction shock: $\Delta(W) = J(W) - W$. This exposure also determines the impact of severance on the agent's continuation value W_t and, thus, on the incentive costs incurred prior to job destruction.

[11] Parameter values used to compute this example are given in Section 8.2.

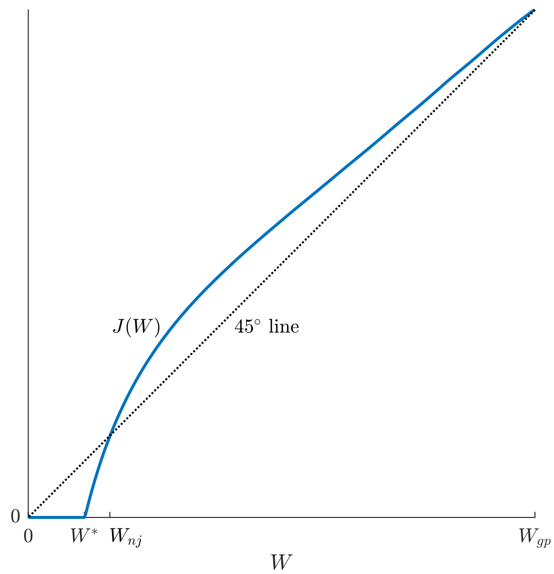

Figure 1: Optimal severance value $J(W)$.

5.1 Reward and punishment at job destruction

Proposition 1 already implies that $\Delta(W) < 0$ for W low enough. In probation $J(W) = 0$, so $\Delta(W) = -W < 0$ for all $W \in (0, W^*]$. Outside of probation, we have the following characterization of the agent's exposure to job destruction risk.

Proposition 2 *In an optimal contract, there exists a unique $W_{nj} \in (W^*, W_{gp})$ such that*

$$\begin{aligned} &\Delta(W) < 0 \text{ for } W \in (0, W_{nj}), \\ &\Delta(W) = 0 \text{ for } W = W_{nj}, \\ &\Delta(W) > 0 \text{ for } W \in (W_{nj}, W_{gp}). \end{aligned}$$

Thus, the agent loses at job destruction if $W_\theta < W_{nj}$ and gains if $W_\theta > W_{nj}$.

Proof In Appendix A.3. ■

The proof of this result boils down to showing that the functions F' and F'_{sep} cross exactly once on $(0, W_{gp})$. Indeed, since F' and F'_{sep} are strictly decreasing, the first-order condition (14) will be met with a negative $\Delta(W) = J(W) - W < 0$ if $F'(W) > F'_{\text{sep}}(W)$, and with a positive $\Delta(W) > 0$ if $F'(W) < F'_{\text{sep}}(W)$.

From the smooth-pasting condition $F'(W_{gp}) = F'_{\text{sep}}(W_{gp})$, clearly, we have that F' and F'_{sep} are equal at W_{gp}. They must cross at least once on $(0, W_{gp})$ because $F'(0) > 0 = F'_{\text{sep}}(0)$ and $F(W) = F_{\text{sep}}(W)$ at both $W = 0$ and $W = W_{gp}$.[12]

[12]Indeed, since $F'(0) > 0 = F'_{\text{sep}}(0)$ and F' and F'_{sep} are continuous, $F'(W) \geq F'_{\text{sep}}(W)$ for all $W \in$

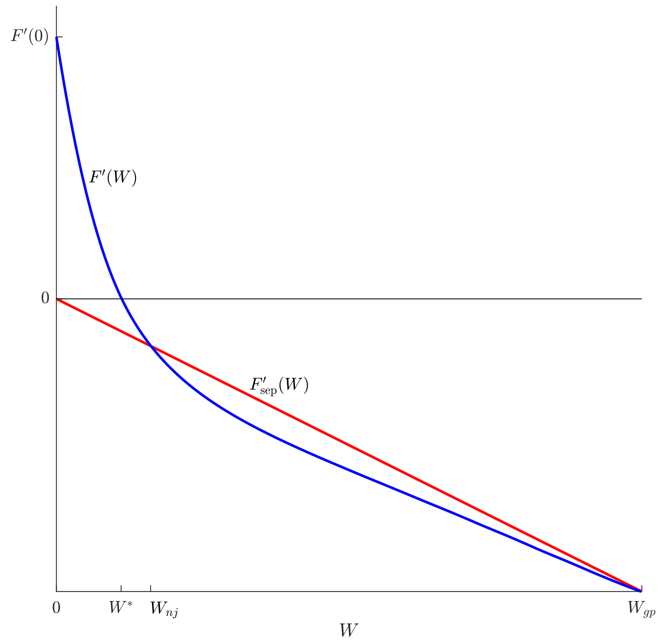

Figure 2: The single crossing between F'_{sep} and F'. For $W > W^*$, the optimal jump of the agent's continuation value at job destruction, $\Delta(W)$, equalizes the two slopes.

The proof that only a single crossing between F' and F'_{sep} exists on $(0, W_{gp})$ would be immediate if F' were convex and F'_{sep} concave on $(0, W_{gp})$. Assumption 1 provides the concavity of F'_{sep}, but the convexity of F' on the whole interval $(0, W_{gp})$ cannot be guaranteed.

However, by differentiating the HJB equation we can show that F' is strictly convex whenever F' is above F'_{sep}, which is enough to eliminate the possibility of a second crossing between F' and F'_{sep} on $(0, W_{gp})$. Indeed, if a second crossing between F' and F'_{sep} were to occur at some $\hat{W} < W_{gp}$, then on the interval $(\hat{W}, W_{gp})$ we would have F' above F'_{sep} and bending upward (because F' is strictly convex), while F'_{sep} is bending (weakly) downwards. This would mean that, on the interval $(\hat{W}, W_{gp})$, F' moves farther and farther away from F'_{sep}, contradicting the smooth-pasting condition $F'(W_{gp}) = F'_{\text{sep}}(W_{gp})$.

Figure 2 shows the single crossing between F' and F'_{sep} in a computed example. The optimal adjustment $\Delta(W)$ is given by the horizontal distance from $F'(W)$ to the line representing F'_{sep}, or to the vertical axis if $F'(W) > 0$.

Assumptions 1 and 2 are convenient sufficient conditions for a single crossing between F' and F'_{sep}, but they are not necessary. In Figure 2, it is easy to see that we can change F'_{sep} slightly to make it slightly convex without substantially changing the shape of F'

$(0, W_{gp})$ would imply $\int_0^{W_{gp}} (F'(W) - F'_{\text{sep}}(W))dW = (F(W) - F_{\text{sep}}(W))|_0^{W_{gp}} > 0$, i.e., we could not have $F(W) = F_{\text{sep}}(W)$ at both $W = 0$ and $W = W_{gp}$.

thus preserving the single crossing between these two curves inside the interval $(0, W_{gp})$. Similarly, a violation of Assumption 2 for W much smaller than W_{nj} will not affect the convexity of F' above W_{nj}, thus preserving the single-crossing property.

5.2 Early and late career: two stages of an optimal contract

We will refer to the interval (W^*, W_{nj}) as the *early career* stage and the interval (W_{nj}, W_{gp}) as the *late career* stage of the optimal contract.

Proposition 2 shows that at arrival of a job destruction shock it is optimal to widen the spread of the continuation value: agents with high W (i.e., in late career) see their continuation value jump upward, while agents with low W (in probation or in early career) experience a drop in their continuation value. In other words, rich agents become richer and poor agents become poorer at the moment of job destruction. This increase in dispersion at job destruction is optimal because the jump Δ has an inverse impact on the drift of W_t prior to job destruction: the threat of a negative Δ pushes W_t upward, and the promise of a positive Δ pushes W_t downward. By suppressing the growth of W_t when W_t is high and boosting the growth of W_t when W_t is low, the optimal severance $J(W_t)$ decreases the risk of hitting either of the two inefficient separation points, 0 and W_{gp}, while the match remains productive.

In particular, a "golden parachute" severance package with $J(W_t) > W_t$ is optimal for agents in late career, where the risk of inefficient retirement at W_{gp} is paramount. In probation, the risk of inefficient termination of the agent at $W = 0$ is the primary concern: combined with the agent's limited liability, this risk implies $J(W_t) = 0$ in this region. Proposition 2 shows that this intuition holds not only when W_t is close to 0 or W_{gp} but everywhere in the support of W_t. The early career stage of an optimal contract is where the balance of these two risks shifts: as W_t increases, the risk of inefficient retirement takes on more and more weight in determining the optimal severance $J(W_t)$.

Further, Lemma 3 in Appendix A.3 implies the following corollary to Proposition 2:

Corollary 1 *In probation and in the early career stage of an optimal contract, the agent's is under-paid, $u(c(W)) \leq J(W) \leq W$, and his continuation value increases in expectation:*

$$\mu(W) > 0 \textit{ for all } W \in (0, W_{nj}).$$

Intuitively, the agent is under-paid in probation and in early career, both with direct compensation and with severance, in order to mitigate the risk of the inefficient endogenous separation at $W = 0$. In these two regions, the expected movement of the agent's continuation value is unambiguously upward, toward the late-career stage of the contract.

5.3 The agent's maximum exposure to job destruction

Proposition 3 *The agent's gain at job destruction, $\Delta(W)$, is zero at both 0 and W_{nj} and has a V-shape on $(0, W_{nj})$:*

$$\Delta'(W) < 0 \text{ for } W \in (0, W^*),$$
$$\Delta'(W) > 0 \text{ for } W \in (W^*, W_{nj}].$$

Thus, W^* is a unique trough point of $\Delta(W)$.

Proof In Appendix A.4. ■

This result shows that the agent stands to lose the most from job destruction at the very beginning of the contract, right in between probation and early career. In probation, the agent gets no compensation of either form. To start receiving compensation, the agent must reach the early-career stage of the contract. The deeper in probation the agent finds himself, thus, the less he stands to lose at job destruction because the expected transition to the early-career stage would take longer should job destruction not occur. In early career, the agent is still at risk of losing value at job destruction, and only upon reaching the late-career stage he stands to gain. The closer he finds himself to the transition point W_{nj}, thus, the smaller his exposure to the job destruction risk.

We should also note that while the agent gains at job destruction in late career, this gain approaches zero at both ends of (W_{nj}, W_{gp}). Thus, the peak of the agent's gain occurs in the interior of the late-career interval.

6 The firm's exposure to job destruction risk

In this section, we examine the firm's exposure to job destruction risk in an optimal contract. We show that severance compensation is used to optimally load the firm's loss at job destruction onto events in which the agent has produced strong performance and the contract has reached the late-career stage.

Proposition 4 *In an optimal contract, the firm's loss of value at job destruction is always strictly positive: $F(W) - F_{\text{sep}}(J(W)) > 0$ for all $W \in (W, W_{gp})$. Further, it is optimal to load the firm's loss at job destruction onto late-career events: the loss is highest if $W_\theta \in (W_{nj}, W_{gp})$.*

Proof In Appendix A.5. ■

To understand this result, it is useful to consider the firm's loss at job destruction under the two severance plans in the examples in Section 3. In Example 1, $J(W)$ is restricted to be 0 for all W. The shape of the firm's loss at job destruction, thus, matches exactly the shape of F, which means the peak loss occurs if $W_\theta = W^*$. In Example 2, $J(W) = W$ for all W. The firm's loss at job destruction, thus, is $F(W) - F_{\text{sep}}(W)$. By the single-crossing property of F' and F_{sep}, this difference has a unique peak at W_{nj}, which means the peak loss occurs if $W_\theta = W_{nj}$.

Proposition 4 shows that in an optimal contract, the firm's loss is shifted even further to the right: the peak loss at job destruction occurs above W_{nj}, i.e., in the late-career stage of the contract, where the risk of inefficient retirement of the agent is most severe. Severance is used to mitigate this risk, i.e., to suppress the drift of W_t and extend the duration of the contract. This extension comes at the cost of a large severance obligation to the agent if job destruction occurs in this stage of the contract.

7 Comparative statics for job destruction risk

In this section, we provide two comparative statics results. We examine how the firm's profit and the back-loading of the agent's compensation depend on the risk of job destruction as measured by the arrival rate λ.

7.1 Sensitivity of profit to job destruction risk

Proposition 5 *$F'(0)$, W_{gp}, and $F(W)$ for any $W \in (0, W_{gp})$ are all strictly decreasing in λ. Further,*

$$\frac{\partial F(W)}{\partial \lambda} = -\mathbb{E}\left[\int_0^\tau e^{-(r+\lambda)t} S(W_t)dt\right] < 0, \tag{17}$$

where $W_0 = W \in (0, W_{gp})$, τ denotes the time of the first exit of W_t from $(0, W_{gp})$, and

$$S(W) \equiv F(W) - F'(W)W - \max_{J \geq 0}\left\{F_{\text{sep}}(J) - F'(W)J\right\} > 0 \quad \forall W \in (0, W_{gp}) \tag{18}$$

represents the value at risk of job destruction, i.e., the value of the option to not separate before job destruction.

Proof In Appendix A.6. ■

It is intuitive that $F(W)$ should be decreasing in λ. Operating the productive technology for a shorter expected duration (i.e., under higher λ) gives the firm a smaller expected profit ex ante. Geometrically, higher λ flattens out the hump-shaped function $F(\cdot)$, which implies that both $F'(0)$ and $-F'(W_{gp})$ are lower. Because W_{gp} is pinned by the smooth-pasting condition in (13), a lower $-F'(W_{gp})$ also shifts W_{gp} to the left. The left end of the domain of F is unaffected by λ because the boundary condition $F(0) = 0$ is independent of λ.

Note that W_{gp} decreasing in λ means that with moral hazard some jobs can be non-viable just because the anticipated job duration is too short. Indeed, consider two jobs in two different occupations or industries with the same productivity and the same continuation value owed to the worker, but with different arrival rates of job destruction. It may be profitable to continue with a worker in the longer-expected-duration job but endogenously terminate (retire) a worker in the job that has lower expected duration, despite both workers being equally productive and owed the same compensation.

In (17), the quantity S can be thought of as representing the value at risk of job destruction or, equivalently, the surplus from the option to not separate. The first term in (18), $F(W) - F'(W)W$, gives the total value of the relationship before job destruction. It is the sum of the firm's profit, $F(W)$, and the agent's value, W, with the slope of the Pareto frontier, $-F'(W)$, representing the shadow price of the agent's value in terms of profits (mapping utils to dollars). The second term represents the relationship's value at separation, where the agent's severance value, J, is selected optimally.[13]

Using S, we can write the HJB equation (11) as:

$$rF(W) = \max_{c,a,Y}\{ra - rc + F'(W)r\big(W - u(c) + h(a)\big) + \frac{1}{2}F''(W)r^2\sigma^2 Y^2\} - \lambda S(W).$$

Mechanically, $S(W)$ captures the direct impact of a change in λ on $F(W)$. The equality in (17) shows that total impact of a change in λ on $F(W_0)$ is equal to the discounted sum of future $S(W_t)$, accounting for the expected position of the process W_t at all $t \geq 0$. The discount factor in (17) contains the probability of job survival till t, $e^{-\lambda t}$, because the firm's option to not separate is conditional on job survival.

The strict inequality in (17), follows from the fact that the option value $S(W)$ is strictly positive for all $W \in (0, W_{gp})$. Additionally, Lemma 6 in Appendix A.6 shows that $S(W)$ approaches zero at both ends of this interval, where endogenous separations indeed occur, and is hump-shaped with a unique peak at W_{nj}.

7.2 Backloading of reward and punishment

We now examine the impact of the risk of job destruction on the timing of the agent's rewards and punishment in an optimal contract.

Proposition 6 *Let $\tilde{\lambda} > \lambda \geq 0$. There exists a unique $W^s \leq \tilde{W}_{gp}$ such that, for all $W \in [0, \tilde{W}_{gp}]$, $\tilde{J}(W) \geq J(W)$ and $\tilde{c}(W) \geq c(W)$ if and only if $W \leq W^s$.*

Proof In Appendix A.7. ■

[13] The outside equality in (16) implies that slope of the Pareto frontier remains unchanged through job destruction if $F'(W_\theta) \leq 0$. In the other case, the agent's limited liability constraint, $J \geq 0$, binds.

This result means that the agent's rewards and punishment, implemented via both compensation and severance, are more backloaded and spread out when the expected duration of the contract is longer. With $\lambda = 0$, the standard moral hazard model of Sannikov (2008) maximally backloads and spreads out the agent's rewards and punishment. Indeed, that model only has endogenous separations with the agent's continuation value either maximally low, $W_\tau = 0$, or maximally high, $W_\tau = W_{gp}$. With exogenous separations, the expected job duration is shorter, and the agent's separation value J_θ is bounded below by 0 and above by W_{gp}. The same bounds apply to the agent's consumption utility flow that the optimal contract delivers to the agent prior to separation.

More precisely, Proposition 6 implies that for any two job destruction arrival rates $\tilde{\lambda} > \lambda$, the dispersion in compensation and severance across agents with different output histories is lower when the expected duration of the contract is shorter. To see this, suppose we start with $\tilde{\lambda} = \lambda$, which obviously means $\tilde{F}'(W) = F'(W)$ for all W. If we increase $\tilde{\lambda}$ slightly above λ, it is easy to see in Figure 2 that the curve $\tilde{F}'$ starts moving away from F' and toward F'_{sep}. Indeed, we know from Proposition 5 that $\tilde{F}'(0)$ decreases toward zero. We know from Proposition 6 that there can be only a single crossing between $\tilde{F}'$ and F', at W^s. Thus, $\tilde{F}'$ rotates counterclockwise as it approaches F'_{sep}, meaning the value of $\tilde{F}'(W)$ goes down at small W and goes up at large W, i.e., $\tilde{F}'$ becomes flatter as $\tilde{\lambda}$ increases.[14] By the first-order conditions (14) and (15), the slope of the first derivative of the profit function determines the dispersion in the agent's compensation and severance. Thus, the dispersion is lower when the arrival rate of job destruction is higher, i.e., the expected contract duration is shorter. Intuitively, job destruction puts a constraint on how backloaded incentives can be.

8 Computation of contract duration and exit probabilities

In this section, we show how to compute additional features of the contract: the conditional expected duration and exit probabilities. We compute a baseline example without job destruction shocks and an example with job destruction to show how job destruction shocks impact the expected duration and exit probabilities of an optimal contract.

[14]This rotation of F' is consistent with the flattening of F shown in Proposition 5 because F is hump-shaped. Indeed, for a hump-shaped function function F to flatten, the positive part of F' must be decreased and the negative part of F' must be increased (made less negative).

8.1 Job duration and exit probability

Let $T(W)$ denote the expected remaining job duration, i.e., the expected amount of time until endogenous separation or exogenous job destruction:

$$T(W) \equiv \mathbb{E}\left[\int_0^{\hat{\tau}} dt\right], \; W_0 = W. \tag{19}$$

Let $P_0(W)$ denote the probability that the continuation value process W_t started at $W_0 = W$ hits 0, i.e., the contract ends with the agent being fired. Similarly, let $P_{gp}(W)$ denote the probability that the contract exits at W_{gp} with the agent being retired. The probability that the contract ends with job destruction, i.e., that a shock θ arrives before W_t reaches either 0 or W_{gp}, will be denoted by $P_{JD}(W)$. Clearly, $P_0(W) + P_{gp}(W) + P_{JD}(W) = 1$ for any $W \in [0, W_{gp}]$.

To compute the expected duration $T(W)$ and exit probabilities $P_0(W)$, $P_{gp}(W)$, and $P_{JD}(W)$ for all $W \in [0, W_{gp}]$, we find an ordinary differential equation (ODE) for each of these functions along with the associated boundary conditions. These ODEs can then be easily solved numerically using policy functions (or drift and volatility) from the optimal contract.

Lemma 1 *Suppose for some constants k, g_0, and g_1, function $g : [0, W_{gp}] \to \mathbb{R}$ satisfies the following ODE:*

$$\lambda g(W) = k + g'(W)\mu(W) + \frac{1}{2}g''(W)\nu(W)^2 \tag{20}$$

with boundary conditions $g(0) = g_0$ and $g(W_{gp}) = g_1$. Then,

$$\begin{aligned}
k = 1,\; g_0 = 0,\; g_1 = 0 \quad &\Longrightarrow \quad g = T,\\
k = 0,\; g_0 = 1,\; g_1 = 0 \quad &\Longrightarrow \quad g = P_0,\\
k = 0,\; g_0 = 0,\; g_1 = 1 \quad &\Longrightarrow \quad g = P_{gp},\\
k = \lambda,\; g_0 = 0,\; g_1 = 0 \quad &\Longrightarrow \quad g = P_{JD}.
\end{aligned}$$

Proof In Appendix A.8. ■

8.2 Computed examples

For the computation of examples, we use the same utility functions as in Sannikov (2008): $h(a) = 0.5a^2 + 0.4a$ and $u(c) = \sqrt{c}$, which gives us $F_0(W) = -W^2$. To isolate on the impact of λ on an optimal contract, we take $F_{\text{sep}}(W) = F_0(W)$ in our examples. To match the standard annualized rate of time preference of 5 percent, we take $r = 0.0488$. Following

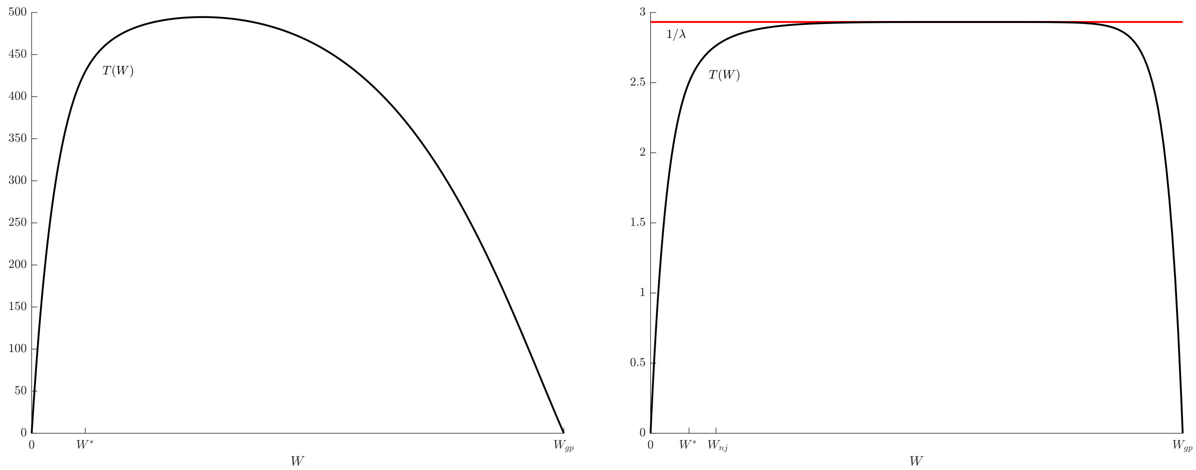

Figure 3: Expected contract duration $T(W)$, in years. Left panel: standard model with no job destruction. Right panel: model with job destruction risk matching $T(W^*) = 2.5$ years.

Sannikov (2008), we take $\sigma = 1$.[15] In the baseline case without job destruction risk, we take $\lambda = 0$. In the example with job destruction risk, we take $\lambda = 0.341$. With this value of λ the expected ex ante duration of the contract is $T(W^*) = 2.5$, which matches the average job duration in the data (Shimer, 2005).

To solve for the value function F and the associated policy functions, we use the forward-shooting procedure described in Appendix A.1. We then use Lemma 1 to compute, for each $W \in [0, W_{gp}]$, the conditional duration $T(W)$ and the conditional exit probabilities $P_0(W)$, $P_{gp}(W)$, and $P_{JD}(W)$.

Figure 3 plots the expected duration function $T(W)$. The left panel shows the baseline example without job destruction. Expected conditional job durations are very long in the baseline case except for W very close to 0 or W_{gp}. The right panel of Figure 3 shows the example with job destruction risk. In this example, the expected time till job destruction is $1/\lambda = 2.93$ years. For W close to W^*, the expected duration is noticeably shorter than $1/\lambda$, which means the contract has a positive chance of endogenous termination. However, in the middle of the domain $[0, W_{gp}]$ duration $T(W)$ is extremely close to $1/\lambda$, which means that once the contract enters this area, the chance of an endogenous exit at either 0 or W_{gp} is practically zero.

Figure 4 plots the conditional exit probabilities $P_0(W)$, $P_{gp}(W)$, and $P_{JD}(W)$. In the left panel, in the baseline case without job destruction, obviously, $P_{JD}(W)$ is zero and

[15]Using Compustat data, Comin and Mulani (2006) estimate the volatility of the growth rate of sales at the firm level to be about 0.25. Although the mapping between the volatility measure they estimate and our volatility parameter σ is not exact, our value of $\sigma = 1$ probably overstates the volatility of shocks to an individual worker's output.

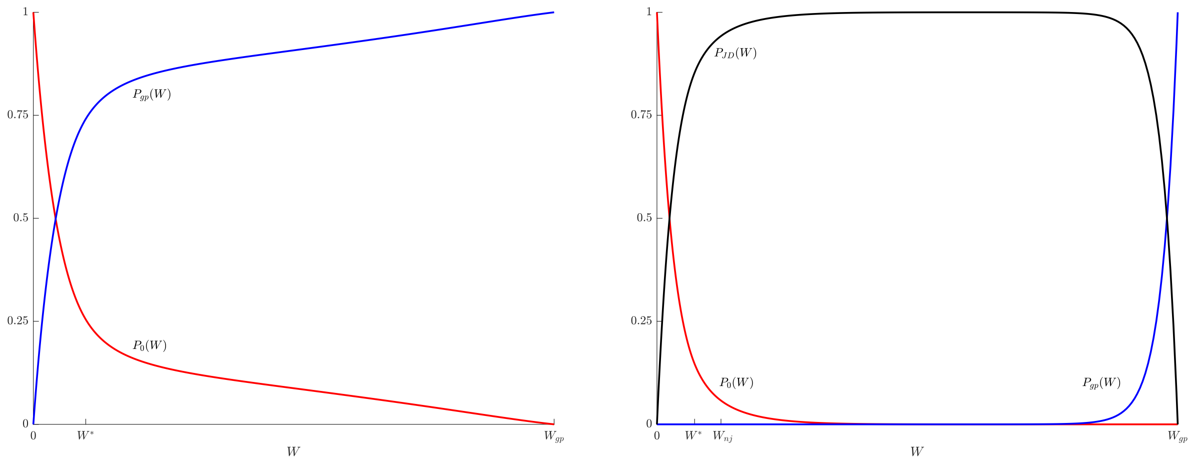

Figure 4: Exit probabilities $P_0(W)$, $P_{gp}(W)$, and $P_{JD}(W)$. Left panel: standard model with no job destruction. Right panel: model with job destruction risk matching $T(W^*) = 2.5$ years.

$P_0(W)$ and $P_{gp}(W)$ are strictly decreasing in the distance from, respectively, 0 and W_{gp}. For W close to 0, $P_0(W)$ is high but $P_{gp}(W)$ is not negligible. Likewise, for W close to W_{gp}, $P_0(W)$ is not insignificant. In the baseline model, thus, an optimal contract shows significant communication (i.e., the probability of transition) between low and high states. The right panel of Figure 4 plots the conditional exit probabilities in the example with job destruction. In this example, there is practically no communication between states W close to 0 and those close to W_{gp}. In the middle of the domain $[0, W_{gp}]$, the probability that the contract ends with arrival of job destruction is practically full.

To help us better understand the dynamics of W_t, Figure 5 plots the drift and volatility functions $\mu(W)$ and $\nu(W)$ for an optimal contract in the two examples. These functions are qualitatively similar in both cases. Starting from W^*, the contract has an initial strong positive drift and sizable volatility. Conditional on not reaching 0, thus, the contract moves up quickly into the middle of the interval $[0, W_{gp}]$. There, however, the movement of W_t "slows down" very significantly. Indeed, in the middle of the interval $[0, W_{gp}]$, as well as at high W, the process W_t has near-zero drift and reduced volatility. This means the contract, upon reaching the middle of its domain, is expected to spend a lot of time in this region. In the baseline case, this feature leads to the long expected contract duration shown in the left panel of Figure 3. In the example with job destruction, this feature implies that job destruction arrives before the contract leaves the middle region, i.e., the chance that the contract eventually ends with an endogenous separation becomes practically zero, as shown in the right panel of Figure 4. Consistent with Corollary 1, in the right panel of Figure 5, $\mu(W) > 0$ for all W in the probation and early career regions, while the region of "slow" contract dynamics is a subset of the late-career region.

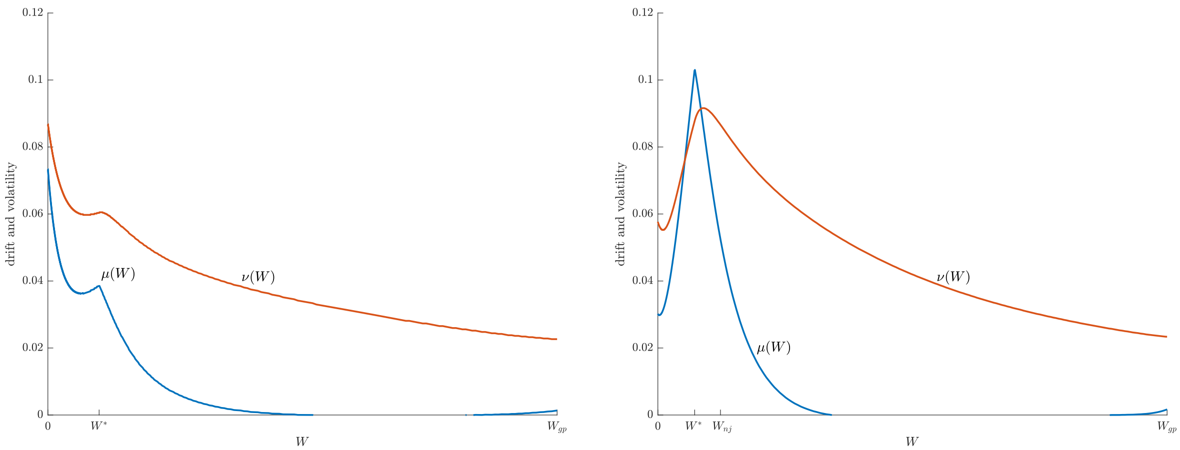

Figure 5: Drift μ and volatility ν under an optimal contract. Left panel: standard model with no job destruction. Right panel: model with job destruction risk matching $T(W^*) = 2.5$ years.

The drift and volatility of the optimal contract in the right panel of Figure 5 also suggest that the realized severance award should be positively correlated with realized tenure. Indeed, an observed short job spell is more likely than a long spell to have ended with an endogenous termination at $W = 0$, or with an exogenous termination in the probation region of the contract, where severance pay is zero. Figure 6 confirms this intuition by plotting realized severance pay against job duration in 100 complete job spells simulated in our model. The positive correlation predicted by our model is broadly consistent with an assumption made in the quantitative macro-labor literature, where severance is commonly assumed to be an increasing function of job tenure, see, e.g., Cozzi and Fella (2016).

We also note that the simulated data in Figure 6 show no instances of agent retirement, i.e., no endogenous terminations at W_{gp}. Clearly, the possibility of job destruction makes endogenous endogenous termination less likely. In particular, as shown in the right panel of Figure 4, with $W_0 = W^*$, the probability of the contract reaching W_{gp} ahead of job destruction is very close to zero. In the next section, we confirm this observation analytically allowing for a sufficiently high rate of job destruction.

9 Endogenous termination when contract duration is short

In this section, we study the limit case as λ goes to infinity, i.e., as the expected job duration becomes short. We examine the survival of the possibility of endogenous separations ahead of exogenous job destruction. We show that endogenous terminations survive even in the limit. In particular, the endogenous termination at 0 survives while the termination at W_{gp} gets eliminated.

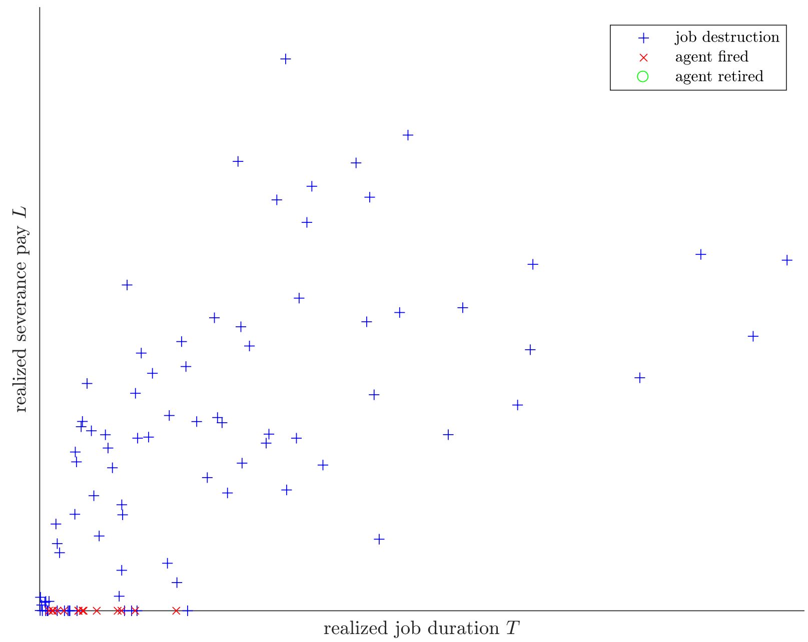

Figure 6: Severance and job duration in 100 simulated complete job spells.

Proposition 7 *Suppose*

$$\kappa < F''_{\text{sep}}(0) < 0. \tag{21}$$

Then,

$$\lim_{\lambda \to \infty} P_0(W^*) > 0 \quad \text{and} \quad \lim_{\lambda \to \infty} P_{gp}(W^*) = 0.$$

Proof In Appendix A.9. ■

The left inequality in (21) ensures that a non-degenerate contract exists. The right inequality in (21) is a convenient technical assumption that may be relaxed.[16]

Intuition for the results of Proposition 7 comes from two effects that an increase in the job destruction arrival rate λ has on endogenous separations. First, clearly, higher λ increases the probability of exogenous separation and hence decreases the probability of endogenous separation. Second, higher λ decreases the agent's ex ante value of the job, thus making $W_0 = W^*$ smaller. Both effects make a transition of W_t from W^* to W_{gp} less likely. The second effect, however, turns out to be strong enough for the endogenous separation at $W = 0$ to survive, even in the limit as λ goes to infinity.

Proposition 7 shows that exogenous job destruction shocks change qualitative properties of an optimal contract. Such shocks not only add the possibility of exogenous separation

[16] We conjecture that the same result holds in many cases with $F''_0(0) = 0$, but we do not offer a proof.

with severance compensation at any t but also, for λ sufficiently high, they eliminate the possibility of endogenous separation at the retirement point W_{gp}. As demonstrated by the numerical example in Section 8.2, the retirement separations are practically eliminated not only in the limit but already at realistic levels of job destruction risk $\lambda > 0$.

10 Testable implications for compensation and tenure

In this section, we show that our model predicts a positive relationship between expected job duration and average compensation. By adjusting the arrival rate of job destruction, λ, we generate optimal contracts with expected job duration ranging from less than 2 months to over 10 years. For each of these contracts, we compute two measures of compensation: the ex ante expected average wage rate and the ex ante expected severance benefit. We show that both are increasing in the ex ante expected job duration.

Denote by $TC(W)$ the expected total wage paid in the remainder of the contract:

$$TC(W) \equiv \mathbb{E}\left[\int_0^{\hat{\tau}} c\left(W_t\right) dt\right], \ W_0 = W.$$[17]

We define the expected average wage in an optimal contract, $\bar{C}$, as the total expected wage in the contract starting at W^* divided by the total expected duration of this contract:

$$\bar{C} \equiv \frac{TC(W^*)}{T(W^*)},$$

where the expected contract duration function $T(W)$ is defined in (19).

To express severance in the units of the firm's profit rather than in utils, let $b = -F_0(J)$ denote the firm's (permanent flow) cost to fund the agent's severance value J.[18] Denote by $B(W)$ the expected cost b of severance to be due to the agent at job destruction, for a contract starting at W:

$$B(W) \equiv \mathbb{E}\left[-F_0(J(W_\theta))\right], \ W_0 = W.$$[19]

The expected severance in an optimal contract is defined as

$$\bar{B} \equiv B(W^*).$$

[17]Similar to Lemma 1, we compute $TC(W)$ by solving an ODE identical to (20) except with the constant k replaced by the policy function $c(W)$. The boundary conditions for $TC(W)$ are $TC(0) = TC(W_{gp}) = 0$.

[18]In the notation of equation (3), $b = rL$. Since we use $F_{\text{sep}} = F_0$ here, $D = 0$ and $V = u(b)$.

[19]We compute $B(W)$ by solving an ODE identical to (20) except with the constant k replaced by $-\lambda F_0(W)$. The boundary conditions for $B(W)$ are $B(0) = 0$ and $B(W_{gp}) = -F_0(W_{gp})$.

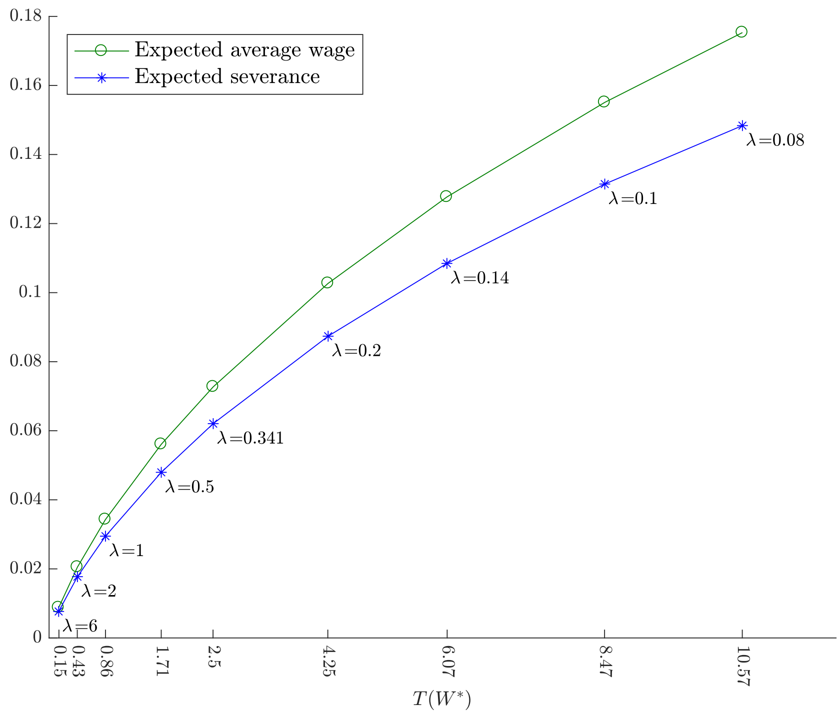

Figure 7: Expected average wage $\bar{C}$ and expected severance $\bar{B}$ in optimal contracts with different expected duration $T(W^*)$.

Figure 7 plots $\bar{C}$ and $\bar{B}$ as a function of $T(W^*)$. It shows that the expected average wage and the expected severance are lower in jobs with shorter expected duration.[20] The relationship between expected job duration and average compensation can be explained through two effects. First, the agent's ex ante value, W^*, is lower when the expected job duration is shorter. Second, recall from Proposition 5 that higher λ flattens out the hump-shaped profit function F. That is, both the increasing segment and the decreasing segment of F are less steep when λ is higher. The first-order condition (15) then implies that the agent's compensation flow, c, weakly decreases with λ: on the increasing segment of F, c is always zero, i.e., is unaffected by λ; on the decreasing segment of F, c is strictly lower when F is less steep. This flattening of F makes the agent's average compensation $\bar{C}$ lower, holding the value delivered to the agent constant. The same intuition applies to the expected severance $\bar{B}$ because the first-order condition (16) implies that direct compensation c and severance compensation $-F_0(J)$ are positively related.

The main testable implication of our model, therefore, is a positive correlation between

[20] Figure 6, by contrast, shows a positive ex post tenure-severance correlation among jobs with the same ex ante expected duration.

expected job duration and average wages and severance. There is a large literature on the individual-level correlation between tenure and wages. This literature, generally using PSID data, e.g., Topel (1991) and Buchinsky et al. (2010), finds a positive wage-tenure profile, which is often attributed to the accumulation of unobservable human capital. This channel is not present in our model, where agent productivity is *iid*. Our model suggests a different channel for generating a positive wage-tenure profile: both tenure and wages can be driven by job destruction risk, which can be heterogeneous across occupations, sectors, localities, or demographic groups.

Our model also generates a positive severance-tenure profile, consistent with the data, as documented in, e.g., Boeri et al. (2017). Although it is not important for our finding concerning the slope of severance against tenure, we note that in our computed example the level of severance constitutes an outsized proportion of average wages. In particular, with the expected job duration of 2.5 years, Figure 7 implies a lump-sum severance equivalent to about 17 years of average wages, which is an order of magnitude higher than what Boeri et al. (2017) find in OECD data.[21] Clearly, our numerical example understates the agent's post-separation outside value, which under the separation profit function F_0 is set equal to zero.

11 Conclusion

Our analysis underscores a simple intuition for why severance pay can be an efficient means of compensation: by deferring a part of the agent's compensation until separation, severance reduces the agent's compensation prior to separation, which can be beneficial for incentives. We show that this intuition holds in the canonical dynamic moral hazard model of Sannikov (2008), where incentive cost increase with the agent's value, but only when this value is sufficiently high. The resulting optimal contract can be split into three stages: probation, where severance is zero; early career, where severance is small; and late career, where a large severance award reduces the incentive costs by mitigating the risk of inefficient retirement of the agent. Our model, thus, provides a rationale awarding high severance to agents in high-rank positions while not awarding any to agents in low-rank, entry-level positions.

Our characterization of an optimal contract shows that the late-career stage is almost absorbing: In late career, incentives are relatively weak, and the agent's continuation value is close to stationary. The contract dynamics slow down, and exogenous job destruction becomes the only viable exit for the contract. In particular, endogenous retirement of

[21] Indeed, the present value of a permanent severance benefit $\bar{B}$ is $r\int_0^\infty e^{-rt}\bar{B}dt = \bar{B}$, while the present value of one year of average wage $\bar{C}$ is $r\int_0^1 e^{-rt}\bar{C}dt = (1-e^{-r})\bar{C}$. At $T(W^*) = 2.5$, we have in Figure 7 $\bar{C} = 0.75$ and $\bar{B} = 0.6$. With $r = 0.0488$, these values imply a ratio $\bar{B}/(1-e^{-r})\bar{C}$ of 16.8.

the agent is almost never observed. This does not mean, however, that the risk of agent retirement becomes unimportant, as it is this risk that shapes the optimal use of severance pay. In an optimal contract, severance pay eliminates this risk almost completely.

We model exogenous separations as a simple Poisson shock that permanently eliminates all productivity in the relationship between the firm and the agent. For our results, it is not essential that the shock eliminates all productivity so long as ending the relationship at the shock's arrival remains optimal. The essential role of severance in our model is to shift the allocation of the agent's value from events with high incentive costs to events with low incentive costs.

Our model delivers testable implications that positively relate the expected job duration with average compensation levels. In jobs less exposed to risk of exogenous destruction, backloading of incentives can be used more extensively, incentive frictions can therefore be resolved more efficiently, resulting with higher profits and compensation levels. In data covering a cross section of jobs, we should thus observe a positive correlation between tenure and compensation.

In the optimal contracting problem we solve, while we use a flexible specification for the agent's continuation value function at separation, we take this function as exogenous. By calibrating V or, equivalently, F_{sep} to the data, quantitative predictions can be obtained on the optimal level and performance sensitivity of severance. Alternatively, an explicit model of a labor market with search frictions can be used to endogenize F_{sep}. Such analysis can provide additional testable implications as well as new theoretical results on the interaction between incentive costs and search costs. Specifically, equilibrium profits and agent value could be non-monotonic in the level of search costs, because more severe search frictions can give rise to longer expected job durations, and, as we show in this paper, incentive costs are lower when expected job durations are longer.

Appendix

A.1 Existence of an optimal solution F

In this section, we verify that the forward-shooting procedure used in Sannikov (2008) to pin down a unique optimal HJB solution F applies with only minor changes to our model, where $F_{\text{sep}} \geq F_0$ and a job destruction shock arrives at rate $\lambda \geq 0$.

It will be convenient to use the following notation

$$U(W) \equiv F(W) - \max_{c\geq 0} \left\{ F'(W)(W - u(c)) - c \right\}, \tag{22}$$

and write the HJB equation (11) as

$$F''(W) = \min_{a \geq 0} \frac{U(W) + \frac{\lambda}{r} S(W) - a - F'(W)h(a)}{\frac{1}{2} r\sigma^2 (h'(a))^2}, \tag{23}$$

where S is defined in (18).

Define

$$\mathbb{B} \equiv \big\{ x : \text{solution } F \text{ to HJB equation (23) started from } (F(0), F'(0)) = (F_{\text{sep}}(0), x) \\ \text{satisfies } F(W) < F_{\text{sep}}(W) \text{ at some } W > 0 \big\}.$$

All $x < 0$ belong to $\mathbb{B}$ because $F'_{\text{sep}}(0) = 0$. It is easy to verify that for a sufficiently large $x > 0$ the solution F stays above F_{sep} at all $W > 0$, i.e., $\mathbb{B}$ is bounded above.

To avoid degenerate cases, in which $W_{gp} = 0$ and the optimal contract calls for immediate separation at $t = 0$, we will follow Sannikov (2008) in assuming that HJB solutions F have sufficient curvature at $W = 0$ to imply that $x = 0$ belongs to $\mathbb{B}$.

Define

$$\kappa \equiv - \max_{a \in \mathcal{A}} \frac{a - F_{\text{sep}}(0)}{\frac{1}{2} r\sigma^2 h'(a)^2} < 0. \tag{24}$$

This constant represents, regardless of the value of λ, the curvature at $W = 0$ of the HJB solution started from the boundary value $F(0) = F_{\text{sep}}(0)$ with the initial slope of $x = F'(0) = 0$. Indeed, with $(F(0), F'(0)) = (F_{\text{sep}}(0), 0)$ we have $U(0) = F_{\text{sep}}(0)$ and $S(0) = 0$, for any λ. Evaluating (23) at $(F(0), F'(0)) = (F_{\text{sep}}(0), 0)$, thus, yields $F''(0) = \kappa$, for any λ.

Assumption 3 *$\kappa < F''_{\text{sep}}(0)$.*

Since $F'_{\text{sep}}(0) = 0$, Assumption 3 implies immediately that $x = 0 \in \mathbb{B}$. By continuity, $\mathbb{B}$ contains strictly positive numbers. Following Lemma 1 in Sannikov (2008), it is easy to show that if the initial slope is positive, $x \geq 0$, the HJB solution F is strictly concave. For any $x \geq 0$, it follows from (23) and $F'' < 0$ that the first derivative F' is bounded below by $\frac{-1}{\gamma}$, where $\gamma = \lim_{a \to 0} h'(a) > 0$, as defined in (1).

To pin down the initial slope $F'(0)$ of an optimal solution, we follow the forward-shooting procedure of Sannikov (2008). Starting with a small $0 < x \in \mathbb{B}$, we gradually increase x until the solution F only touches F_{sep} at some point, where smooth pasting conditions (13) hold. Define $\bar{x} \equiv \sup_x \mathbb{B}$. Since $\mathbb{B}$ is bounded above and contains some $x > 0$, $\bar{x}$ is finite and strictly positive. Recall W^*_{gp} defined in Assumption 2 as the solution to $F'_{\text{sep}}(W) = -1/\gamma$.

Lemma 2 *If $F'(0) = \bar{x}$, then $F(W) \geq F_{\text{sep}}(W)$ for all $W > 0$, and there exists $W_{gp} \in (0, W^*_{gp}]$ at which the smooth-pasting conditions (13) are met.*

Proof First, if $F'(0) = \bar{x}$ and $F(W) < F_{\text{sep}}(W)$ at some $W > 0$, then, by continuity, F still goes under F_{sep} when $F'(0)$ is slightly above $\bar{x}$. This contradicts the fact that $\bar{x} = \sup_x \mathbb{B}$.

Second, there exists some $W_{gp} > 0$ such that $F(W_{gp}) = F_{\text{sep}}(W_{gp})$. To prove this, consider the sequence $x_n \equiv \bar{x} - \frac{1}{n}$ approaching $\bar{x}$ from below. Since $x_n \in \mathbb{B}$, the solution curve F_n starting with $F_n'(0) = x_n$ goes under F_{sep}. Denote the first crossing point by $W_n \equiv \min\{W > 0 : F_n(W) = F_{\text{sep}}(W)\}$. Since $x_n < x_{n+1}$ and $F_n < F_{n+1}$, we have $0 < W_n < W_{n+1}$. We have $W_n < W_{gp}^*$ because $F_{\text{sep}}'(W_n) \geq F_n'(W_n) > \frac{-1}{\gamma} = F_{\text{sep}}'(W_{gp}^*)$. Therefore, being both increasing and bounded, the sequence $\{W_n\}$ has a limit $W_{gp} \equiv \lim_{n\to\infty} W_n \in (0, W_{gp}^*]$. We have

$$F(W_{gp}) = \lim_{n\to\infty} F_n(W_n) = \lim_{n\to\infty} F_{\text{sep}}(W_n) = F_{\text{sep}}(W_{gp}),$$

where the first equality follows from the continuity of $F(W)$ in x and W, and the second equality follows from $F_n(W_n) = F_{\text{sep}}(W_n) \forall n$, and the third from the continuity of F_{sep}.

Third, $F'(W_{gp}) = F_{\text{sep}}'(W_{gp})$ follows from $F(W_{gp}) = F_{\text{sep}}(W_{gp})$ and $F(W) \geq F_{\text{sep}}(W)$ for all $W > 0$.

■

The HJB solution F with initial slope $F'(0) = \bar{x}$ provides policy functions from which an optimal contract is constructed and verified, as in Sannikov (2008). Because $\bar{x} > 0$ and $W_{gp} > 0$, the optimal contract is non-degenerate for any $\lambda \geq 0$. In sum, Assumption 3 guarantees the existence of a non-degenerate optimal contract, even if λ is high.

We maintain Assumption 3 in this paper. However, we note in (23) that higher λ reduces the curvature of the solution F, i.e., makes $F''(W)$ less negative. In the forward-shooting procedure, therefore, higher λ makes it harder for the solution curve F to return to F_{sep}. If Assumption 3 does not hold, it is possible, with sufficiently high $\lambda > 0$, that $0 \notin \mathbb{B}$, i.e., all solutions with initial slope $x \geq 0$ stay above $F_{\text{sep}}(W)$ for all $W > 0$. In this case, the optimal contract is degenerate, i.e., $W_{gp} = 0$ and immediate separation is optimal.

A.2 Proof of Proposition 1

For $W \in (0, W^*]$, the conclusion follows from (14), (15), and $F'(W) \geq 0$. For $W \in (W^*, W_{gp})$, differentiation of (16) yields

$$F_{\text{sep}}''(J(W))J'(W) = \frac{1}{u''(c(W))}c'(W) = F''(W). \tag{25}$$

The strict concavity of u, F, and F_{sep} implies $J'(W) > 0$ and $c'(W) > 0$.
Finally, for any $W < W_{gp}$, we have $F_{\text{sep}}'(J(W)) = F'(W) > F'(W_{gp}) = F_{\text{sep}}'(W_{gp})$, where

the first equality comes from (16), the inequality follows from the strict concavity of F, and the second equality comes from (13). It then follows from the strict concavity of F_{sep} that $J(W) < W_{gp}$.

QED

A.3 Proof of Proposition 2

We start by proving two auxiliary lemmas.

Lemma 3 *For any $W \in (0, W_{gp})$, $F'(W) \geq F'_{\text{sep}}(W)$ implies $u(c(W)) \leq J(W) \leq W$, $\mu(W) > 0$, and $F'''(W) > 0$.*

Proof If $F'(W) \geq 0$, then, by Proposition 1, $u(c(W)) = J(W) = 0$, which means $u(c(W)) = J(W) < W$. If $F'(W) < 0$, we have

$$F'_{\text{sep}}(u(c(W))) \geq F'_0(u(c(W))) = \frac{-1}{u'(c(W))} = F'_{\text{sep}}(J(W)) = F'(W) \geq F'_{\text{sep}}(W),$$

where the first inequality comes from Assumption 2. The three equalities follow from the definition of F_0 and the first-order conditions (16). Since F'_{sep} is decreasing, we thus have $u(c(W)) \leq J(W) \leq W$. Thus, for any $W \in (0, W_{gp})$, $F'(W) \geq F'_{\text{sep}}(W)$ implies

$$\mu(W) = r(W - u(c(W))) + \lambda(W - J(W)) + rh(a(W)) > 0, \tag{26}$$

where the strict inequality follows from $a(W) > 0$ for all $W \in (0, W_{gp})$, which is implied by (1).

Differentiating the HJB equation (11) and canceling out like terms, we obtain

$$0 = F''(W)\mu(W) + \frac{1}{2}F'''(W)\nu(W)^2, \tag{27}$$

where $\nu(W) > 0$, as in (10). It now follows from (26) and $F'' < 0$ that $F'''(W)\nu(W)^2 > 0$. ■

Lemma 4 *For any $W \in (0, W_{gp})$, $F'(W) = F'_{\text{sep}}(W)$ implies $F''(W) < F''_{\text{sep}}(W)$.*

Proof By contradiction, suppose for some $W_1 \in (0, W_{gp})$

$$F'(W_1) = F'_{\text{sep}}(W_1) \text{ and } F''(W_1) \geq F''_{\text{sep}}(W_1). \tag{28}$$

First, we show that there exists $\epsilon > 0$ such that

$$F'(W) > F'_{\text{sep}}(W) \text{ for all } W \in (W_1, W_1 + \epsilon]. \tag{29}$$

The conclusion is obvious if $F''(W_1) > F''_{\text{sep}}(W_1)$. If $F''(W_1) = F''_{\text{sep}}(W_1)$, the conclusion follows from $F'''(W_1) > 0 \geq F'''_{\text{sep}}(W_1)$, where the first inequality follows from Lemma 3 and the second from Assumption 1.

Define now $\bar{W} \equiv \min\{W > W_1 : F'(W) = F'_{\text{sep}}(W)\}$. The smooth-pasting condition (13) implies that $\bar{W}$ exists and $\bar{W} \leq W_{gp}$. Equation (29) implies $\bar{W} > W_1 + \epsilon$. From definition of $\bar{W}$, $F'(W) > F'_{\text{sep}}(W)$ for $W \in (W_1, \bar{W})$. Then, Lemma 3 implies that F' is strictly convex on $(W_1, \bar{W})$. We thus have

$$F'(\bar{W}) > F'(W_1) + F''(W_1)(\bar{W} - W_1) \geq F'_{\text{sep}}(W_1) + F''_{\text{sep}}(W_1)(\bar{W} - W_1) \geq F'_{\text{sep}}(\bar{W}),$$

where the first inequality follows from the strict convexity of F' on $(W_1, \bar{W})$, the second from (28), and the third from the concavity of F'_{sep}, i.e., Assumption 1. This strict inequality contradicts $F'(\bar{W}) = F'_{\text{sep}}(\bar{W})$. ■

The next lemma shows that F' and F'_{sep} cross once on $(0, W_{gp})$.

Lemma 5 *There exists a unique $W_{nj} \in (W^*, W_{gp})$ such that*

$$\begin{aligned} &F'(W) > F'_{\text{sep}}(W) \text{ for } W \in (0, W_{nj}), \\ &F'(W) = F'_{\text{sep}}(W) \text{ for } W = W_{nj}, \\ &F'(W) < F'_{\text{sep}}(W) \text{ for } W \in (W_{nj}, W_{gp}). \end{aligned}$$

Proof The proof proceeds in four steps.

First, we define $W_{nj} \equiv \min\{W : F'(W) = F'_{\text{sep}}(W)\}$. It follows from $F'(0) > 0 = F'_{\text{sep}}(0)$ and $F'(W_{gp}) = F'_{\text{sep}}(W_{gp})$ that $0 < W_{nj} \leq W_{gp}$. We show $W_{nj} < W_{gp}$. By contradiction, suppose $W_{nj} = W_{gp}$. Then, $F'(W) > F'_{\text{sep}}(W)$ for all $0 < W < W_{gp}$. But then $F(W_{gp}) = F(0) + \int_0^{W_{gp}} F'(W)dW > F_{\text{sep}}(0) + \int_0^{W_{gp}} F'_{\text{sep}}(W)dW = F_{\text{sep}}(W_{gp})$, which contradicts (13).

Second, Lemma 4 implies $F''(W_{nj}) < F''_{\text{sep}}(W_{nj})$. Thus, there exists $\epsilon > 0$ such that

$$F'(W) < F'_{\text{sep}}(W) \text{ for all } W \in (W_{nj}, W_{nj} + \epsilon]. \tag{30}$$

Third, we will show $F'(W) < F'_{\text{sep}}(W)$ for all $W \in (W_{nj}, W_{gp})$. By contradiction, suppose it is not true. By (30), F' must be equal to F'_{sep} at some point on $(W_{nj}+\epsilon, W_{gp})$. Denote the smallest such a point by $\hat{W} \equiv \min\{W \in (W_{nj}, W_{gp}) : F'(W) = F'_{\text{sep}}(W)\}$. It now follows from $F'(W) < F'_{\text{sep}}(W)$ for $W \in (W_{nj}, \hat{W})$ and $F'(\hat{W}) = F'_{\text{sep}}(\hat{W})$ that $F''(\hat{W}) \geq F''_{\text{sep}}(\hat{W})$. But Lemma 4 implies $F''(\hat{W}) < F''_{\text{sep}}(\hat{W})$, a contradiction.

Fourth, $W_{nj} > W^*$ follows from $F'(W^*) = 0 > F'_{\text{sep}}(W_{nj}) = F'(W_{nj})$ and F' is continuous and decreasing. ■

The proof of Proposition 2 follows from Lemma 5 and the observation that the sign of $\Delta(W)$ is the same as the sign of $F'_{\text{sep}}(W) - F'(W)$.

QED

A.4 Proof of Proposition 3

For $W \in (0, W^*)$, we have $F'(W) > 0$, which implies $J(W) = 0$. Thus, clearly, $\Delta'(W) = J'(W) - 1 = -1 < 0$ for all $W \in (0, W^*)$.

For $W \in (W^*, W_{gp}]$, the outside equality in (25) shows $J'(W) = \frac{F''(W)}{F''_{\text{sep}}(J(W))}$. Thus, $\Delta'(W) = J'(W) - 1 > 0$ if and only if $F''(W) < F''_{\text{sep}}(J(W))$, which does hold for all $W \in (W^*, W_{nj}]$ because

$$F''(W) \leq F''(W_{nj}) < F''_{\text{sep}}(W_{nj}) \leq F''_{\text{sep}}(J(W)).$$

The first inequality follows from the strict convexity of F' on $(0, W_{nj})$, which is implied by Lemma 3. The second inequality follows from Lemma 4, and the third inequality follows from $F'''_{\text{sep}} \leq 0$ and $J(W) \leq W \leq W_{nj}$.

QED

A.5 Proof of Proposition 4

At $W = 0$ and $W = W_{gp}$, we have $J(W) = W$ and $F(W) = F_{\text{sep}}(W)$. Then, clearly, $F(W) = F(J(W)) = F_{\text{sep}}(J(W))$ at either of these two points.

To show that $F(W) - F_{\text{sep}}(J(W)) > 0$ for all $W \in (0, W_{nj}]$ and that $\operatorname{argmax}_W F(W) - F_{\text{sep}}(J(W)) > W_{nj}$, we will show that $F(W) - F_{\text{sep}}(J(W))$ is strictly increasing on $(0, W^*)$ and on $(W^*, W_{nj}]$. On $(0, W^*)$, the conclusion follows from $J(W) = 0$ and $F'(W) > 0$ for all $W \in (0, W^*)$. On $(W^*, W_{nj}]$, the conclusion follows by differentiating $F(W) - F_{\text{sep}}(J(W))$ and using (16):

$$\begin{aligned}
\frac{d}{dW}\Big(F(W) - F_{\text{sep}}(J(W))\Big) &= F'(W) - F'_{\text{sep}}(J(W))J'(W) \\
&= F'(W)(1 - J'(W)) \\
&= -F'(W)\Delta'(W) \\
&> 0,
\end{aligned}$$

where the inequality follows because $F'(W) < 0$ and, by Proposition 3, $\Delta'(W) > 0$ for all $W \in (W^*, W_{nj}]$.

It remains to be shown that $F(W) - F_{\text{sep}}(J(W)) > 0$ for all $W \in (W_{nj}, W_{gp})$. On this interval, we have $J(W) > W$. Therefore:

$$F(W) - F_{\text{sep}}(J(W)) > F(W) - F_{\text{sep}}(W) = \int_W^{W_{gp}} (F'_{\text{sep}}(x) - F'(x))dx > 0,$$

where the first inequality follows from $F'_{\text{sep}} < 0$ and the second from the boundary condition $F(W_{gp}) = F_{\text{sep}}(W_{gp})$ and Lemma 5.

QED

A.6 Proof of Proposition 5

We start with an axillary lemma characterizing S.

Lemma 6 *$S(0) = S(W_{gp}) = 0$. $S'(W) > 0$ for all $W \in (0, W_{nj})$, $S'(W_{nj}) = 0$, and $S'(W) < 0$ for all $W \in (W_{nj}, W_{gp})$. Thus, W_{nj} is a unique peak point of $S(W)$.*[22]

Proof $S(0) = S(W_{gp}) = 0$ follows directly from the boundary conditions $F(0) = F_{\text{sep}}(0)$, $F(W_{gp}) = F_{\text{sep}}(W_{gp})$. Differentiating $S(W)$ we have

$$\begin{aligned} S'(W) &= F'(W) - F''(W)(W - J(W)) - F'(W) \\ &= F''(W)\,(J(W) - W) \\ &= F''(W)\Delta(W). \end{aligned} \tag{31}$$

Because $F'' < 0$, Δ and S' are of opposite signs. The conclusion follows now from Proposition 2. ■

We now show (17). Differentiating the HJB equation (11) with respect to λ and using $S(W) = F(W) - \max_{J \geq 0} \{F'(W)(W - J) + F_{\text{sep}}(J)\}$, we have

$$(r + \lambda)\frac{\partial F(W)}{\partial \lambda} = \frac{\partial F'(W)}{\partial \lambda}\,(rW - r(u(c) - h(a)) - \lambda(J - W)) + \frac{1}{2}\frac{\partial F''(W)}{\partial \lambda} r^2\sigma^2 Y^2 - S(W),$$

where the controls c, a, and Y are optimal for the value function F. Denoting $\frac{\partial F(W)}{\partial \lambda}$ by $G(W)$ and collecting terms, we thus have the following second-order differential equation:

$$(r + \lambda)G(W) = G'(W)\,(rW - r(u(c) - h(a)) - \lambda(J - W)) + \frac{1}{2}G''(W)r^2\sigma^2 Y^2 - S(W),$$

which, using (9) and (10), we can write simply as

$$(r + \lambda)G(W) = -S(W) + G'(W)\mu(W) + \frac{1}{2}G''(W)\nu(W)^2. \tag{32}$$

At $W = 0$, we have $G(0) = \frac{\partial F(0)}{\partial \lambda} = 0$ because $F(0) = 0$ for all λ. To obtain a boundary condition for G at $W = W_{gp}$, differentiate the boundary condition $F(W_{gp}) = F_{\text{sep}}(W_{gp})$ totally with respect to λ:

$$\frac{\partial F(W_{gp})}{\partial \lambda} + F'(W_{gp})\frac{dW_{gp}}{d\lambda} = F'_{\text{sep}}(W_{gp})\frac{dW_{gp}}{d\lambda},$$

which gives us

$$G(W_{gp}) = \frac{\partial F(W_{gp})}{\partial \lambda} = (-F'(W_{gp}) + F'_{\text{sep}}(W_{gp}))\frac{dW_{gp}}{d\lambda} = 0,$$

[22] Note that $S(W)$ does not have to be concave.

where the last inequality uses the smooth-pasting condition $F'(W_{gp}) = F'_{\text{sep}}(W_{gp})$.

To derive the equality in (17), let

$$H_t \equiv -\int_0^t e^{-(r+\lambda)s} S(W_s)ds + e^{-(r+\lambda)t} G(W_t).$$

We have

$$dH_t = -e^{-(r+\lambda)t} S(W_t)dt - (r+\lambda)e^{-(r+\lambda)t} G(W_t)dt + e^{-(r+\lambda)t} dG(W_t),$$

and so, applying Ito's lemma to compute $dG(W_t)$, we have

$$\begin{aligned} e^{(r+\lambda)t} dH_t = & - S(W_t)dt - (r+\lambda)G(W_t)dt + \left(G'(W_t)\mu(W_t) + \frac{1}{2}\nu(W_t)^2 G''(W_t) \right) dt \\ & + G'(W_t)\nu(W_t)dZ_t. \end{aligned}$$

The dt terms sum up to zero by (32), and $\mathbb{E}[H_t]$ is bounded, i.e., H_t is a martingale. Thus, with τ being a stopping time, we have

$$G(W_0) = H_0 = \mathbb{E}[H_\tau] = \mathbb{E}\left[-\int_0^\tau e^{-(r+\lambda)t} S(W_t)dt + e^{-(r+\lambda)\tau} G(W_\tau) \right].$$

Using the boundary conditions $G(W_\tau) = G(0) = G(W_{gp}) = 0$, we obtain

$$\frac{\partial F(W_0)}{\partial \lambda} = G(W_0) = \mathbb{E}\left[-\int_0^\tau e^{-(r+\lambda)t} S(W_t)dt \right].$$

The strict inequality in (17), i.e.,

$$\frac{\partial F(W_0)}{\partial \lambda} = G(W_0) < 0 \text{ for all } W_0 \in (0, W_{gp}),$$

follows now from $S(W) > 0$ for all $W \in (0, W_{gp})$, which is impled by Lemma 6.

Next, we show that W_{gp} is strictly decreasing in λ. Let $\tilde{\lambda} > \lambda$, and denote by $\tilde{F}$ and $\tilde{W}_{gp}$, respectively, the firm's value function and the upper endogenous separation point obtained in the optimal contract with the job destruction shock arrival rate $\tilde{\lambda}$. We show $\tilde{W}_{gp} < W_{gp}$. Clearly, $\tilde{W}_{gp} \le W_{gp}$ because the first step implies that $F > \tilde{F}$ on $(0, \tilde{W}_{gp})$. To show $\tilde{W}_{gp} < W_{gp}$, suppose by contradiction $W_{gp} = \tilde{W}_{gp}$. This implies

$$F(W_{gp}) = \tilde{F}(W_{gp}) = F_{\text{sep}}(W_{gp}) \text{ and } F'(W_{gp}) = \tilde{F}'(W_{gp}) = F'_{\text{sep}}(W_{gp}). \tag{33}$$

First-order conditions for c and a, and HJB equation (23) immediately imply $c(W_{gp}) = \tilde{c}(W_{gp})$, $a(W_{gp}) = \tilde{a}(W_{gp})$, $J(W_{gp}) = \tilde{J}(W_{gp})$, and $F''(W_{gp}) = \tilde{F}''(W_{gp})$.

1. Differentiating the HJB equation (23) yields

$$F'''(W) = \frac{(W - u(c)) + \frac{\lambda}{r}(W - J(W)) + h(a)}{\frac{1}{2}r\sigma^2(h'(a))^2}(-F''(W)), \tag{34}$$

which implies $F'''(W_{gp}) = \tilde{F}'''(W_{gp})$, because $c(W_{gp}) = \tilde{c}(W_{gp})$, $a(W_{gp}) = \tilde{a}(W_{gp})$, $\lambda(W_{gp} - J(W_{gp})) = \tilde{\lambda}(W_{gp} - \tilde{J}(W_{gp})) = 0$, and $F''(W_{gp}) = \tilde{F}''(W_{gp})$.

2. We show that $a'(W) = \tilde{a}'(W)$ at $W = W_{gp}$. The first-order condition for a is

$$(1 + F'(W)h'(a))h'(a) + \left(U(W) + \frac{\lambda}{r}S(W) - a - F'(W)h(a)\right)2h''(a) = 0.$$

$U'(W_{gp}) = -F''(W_{gp})(W_{gp} - u(c(W_{gp}))) = -\tilde{F}''(W_{gp})(W_{gp} - u(\tilde{c}(W_{gp}))) = \tilde{U}'(W_{gp})$. From (31) and $\Delta(W_{gp}) = 0$, we have $S'(W_{gp}) = F''(W_{gp})\Delta(W_{gp}) = 0$. By the implicit function theorem, thus, $\frac{da}{dW}|_{W=W_{gp}}$ is independent of λ.

3. We now show that at $W = W_{gp}$, $\frac{d(J(W))}{dW} = \frac{F''(W_{gp})}{F''_{\text{sep}}(W_{gp})} < 1$. By contradiction, suppose $F''(W_{gp}) \le F''_{\text{sep}}(W_{gp}) < 0$. Because $F'''(W_{gp}) > 0 \ge F'''_{\text{sep}}(W_{gp})$, the Taylor expansion of $F'(W_{gp} - \epsilon)$ shows that $F'(W_{gp} - \epsilon) > F'_{\text{sep}}(W_{gp} - \epsilon)$ for all small $\epsilon > 0$, which contradicts the fact that $F(W_{gp} - \epsilon) \ge F_{\text{sep}}(W_{gp} - \epsilon)$ for small $\epsilon > 0$.

4. We show $F''''(W_{gp}) < \tilde{F}''''(W_{gp})$. We have

$$\begin{aligned} F''''(W) \quad = \quad & \frac{W - u(c) + \frac{\lambda}{r}(W - J(W)) + h(a)}{\frac{1}{2}r\sigma^2(h'(a))^2}(-F'''(W)) \\ & + \frac{\partial}{\partial a}\left(\frac{W - u(c) + \frac{\lambda}{r}(W - J(W)) + h(a)}{\frac{1}{2}r\sigma^2(h'(a))^2}\right)a'(W)(-F''(W)) \\ & + \frac{1}{\frac{1}{2}r\sigma^2(h'(a))^2}\left(\frac{d(W - u(c))}{dW} + \frac{\lambda}{r}\frac{d(W - J(W))}{dW}\right)(-F''(W)). \end{aligned}$$

Because $\frac{d(W-J(W))}{dW} = 1 - \frac{dJ(W)}{dW} > 0$ and $\lambda < \tilde{\lambda}$, we have $F''''(W_{gp}) < \tilde{F}''''(W_{gp})$. The Taylor expansion now shows that $F(W) < \tilde{F}(W)$ holds in a neighborhood of W_{gp}. This contradicts the fact that $\tilde{F}(W_{gp} - \epsilon)$ stays below $F(W_{gp} - \epsilon)$ for small $\epsilon > 0$.

Finally, we show that if $\tilde{\lambda} > \lambda$, then $F'(0) > \tilde{F}'(0)$. That $F'(0) \ge \tilde{F}'(0)$ follows from $F(0) = \tilde{F}(0) = 0$ and $F > \tilde{F}$ on $(0, \tilde{W}_{gp})$. To show $F'(0) > \tilde{F}'(0)$, suppose by contradiction $F'(0) = \tilde{F}'(0)$. Using the same proof as in the second step, we can show $F''(0) = \tilde{F}''(0)$, $F'''(0) = \tilde{F}'''(0)$, and $F''''(0) < \tilde{F}''''(0)$. They imply that $F(W) < \tilde{F}(W)$ for W near 0, which contradicts the fact that $F > \tilde{F}$ on $(0, \tilde{W}_{gp})$.

QED

A.7 Proof of Proposition 6

By the first-order conditions (15) and (14), it is sufficient to show that $\tilde{F}'$ and F' cross at most once. If they do not cross, then $\tilde{F}'$ is always below F' because, by Proposition 5, $\tilde{F}'(0) < F'(0)$. Then $J(W) \leq \tilde{J}(W)$ and $c(W) \leq \tilde{c}(W)$ for all $W \in [0, \tilde{W}_{gp}]$, and we define $W^s \equiv \tilde{W}_{gp}$. If they cross, we show below that they cross only once. The proof consists of two main steps.

First, we show that if $\tilde{F}'(W) \geq F'(W)$ at some $W \geq W_{nj}$, then $\tilde{F}' > F'$ for all $\tilde{W} > W$. By contradiction, suppose either $\tilde{F}'(W_2) < F'(W_2)$ or $\tilde{F}'(W_2) = F'(W_2)$ at some $W_2 > W \geq W_{nj}$.

1. If $\tilde{F}'(W_2) < F'(W_2)$, then define

$$W_1 \equiv \inf\{\tilde{W} < W_2 : \tilde{F}'(\tilde{W}) < F'(\tilde{W})\},$$
$$W_3 \equiv \sup\{\tilde{W} > W_2 : \tilde{F}'(\tilde{W}) < F'(\tilde{W})\}.$$

$W_3 < \tilde{W}_{gp}$ because $\tilde{F}'(\tilde{W}_{gp}) = F'_{\text{sep}}(\tilde{W}_{gp}) > F'(\tilde{W}_{gp})$. $\tilde{F}' = F'$ at both W_1 and W_3. It follows from $\tilde{F}''(W_1) \leq F''(W_1)$ and

$$\tilde{F}''(W_1) = \frac{\tilde{U}(W_1) + \frac{\tilde{\lambda}}{r}\tilde{S}(W_1) - a - \tilde{F}'(W_1)h(a)}{\frac{1}{2}r\sigma^2(h'(a))^2}$$
$$F''(W_1) = \frac{U(W_1) + \frac{\lambda}{r}S(W_1) - a - F'(W_1)h(a)}{\frac{1}{2}r\sigma^2(h'(a))^2}$$

that $\tilde{U}(W_1) + \frac{\tilde{\lambda}}{r}\tilde{S}(W_1) \leq U(W_1) + \frac{\lambda}{r}S(W_1)$. Therefore,

$$\begin{aligned}
\tilde{U}(W_3) + \frac{\tilde{\lambda}}{r}\tilde{S}(W_3) &= \tilde{U}(W_1) + \frac{\tilde{\lambda}}{r}\tilde{S}(W_1) + (\tilde{U}(W_3) - \tilde{U}(W_1)) + \frac{\tilde{\lambda}}{r}(\tilde{S}(W_3) - \tilde{S}(W_1)) \\
&\leq U(W_1) + \frac{\lambda}{r}S(W_1) + (\tilde{U}(W_3) - \tilde{U}(W_1)) + \frac{\tilde{\lambda}}{r}(\tilde{S}(W_3) - \tilde{S}(W_1)) \\
&< U(W_1) + \frac{\lambda}{r}S(W_1) + (U(W_3) - U(W_1)) + \frac{\tilde{\lambda}}{r}(S(W_3) - S(W_1)) \\
&< U(W_1) + \frac{\lambda}{r}S(W_1) + (U(W_3) - U(W_1)) + \frac{\lambda}{r}(S(W_3) - S(W_1)) \\
&= U(W_3) + \frac{\lambda}{r}S(W_3),
\end{aligned} \tag{35}$$

where the first inequality follows from $\tilde{U}(W_1) + \frac{\tilde{\lambda}}{r}\tilde{S}(W_1) \leq U(W_1) + \frac{\lambda}{r}S(W_1)$, the second from

$$
\begin{aligned}
(\tilde{U}(W_3) - \tilde{U}(W_1)) - (U(W_3) - U(W_1)) &= (\tilde{U}(W_3) - U(W_3)) - (\tilde{U}(W_1) - U(W_1)) \\
&= (\tilde{F}(W_3) - F(W_3)) - (\tilde{F}(W_1) - F(W_1)) \\
&= \int_{W_1}^{W_3} (\tilde{F}'(W) - F'(W))dW < 0
\end{aligned}
$$

and

$$
\begin{aligned}
(\tilde{S}(W_3) - \tilde{S}(W_1)) - (S(W_3) - S(W_1)) &= (\tilde{S}(W_3) - S(W_3)) - (\tilde{S}(W_1) - S(W_1)) \\
&= (\tilde{F}(W_3) - F(W_3)) - (\tilde{F}(W_1) - F(W_1)) \\
&= \int_{W_1}^{W_3} (\tilde{F}'(W) - F'(W))dW < 0,
\end{aligned}
$$

and the third inequality from $W_1 \geq W_{nj}$, $S(W_3) - S(W_1) = \int_{W_1}^{W_3} F''(W)(J(W) - W)dW < 0$, and $\tilde{\lambda} > \lambda$. But $\tilde{F}''(W_3) \geq F''(W_3)$ implies $\tilde{U}(W_3) + \frac{\tilde{\lambda}}{r}\tilde{S}(W_3) \geq U(W_3) + \frac{\lambda}{r}S(W_3)$, contradicting (35).

2. Suppose $\tilde{F}'(W_2) = F'(W_2)$ at some $W_2 > W \geq W_{nj}$. Since part 1 shows $\tilde{F}' \geq F'$ everywhere above W, we have $\tilde{F}''(W_2) = F''(W_2)$, which implies $\tilde{U}(W_2) + \frac{\tilde{\lambda}}{r}\tilde{S}(W_2) = U(W_2) + \frac{\lambda}{r}S(W_2)$ and $a(W_2) = \tilde{a}(W_2)$. It follows from (34), $\tilde{\lambda} > \lambda$, and $W_2 < J(W_2) = \tilde{J}(W_2)$ that $\tilde{F}'''(W_2) < F'''(W_2)$. This implies $\tilde{F}'(W) < F'(W)$ for W near W_2, contradicting part 1.

Second, define $\underline{W}$ as the first crossing point: $\underline{W} \equiv \min\{W : \tilde{F}'(W) = F'(W)\}$. We show that the curves $\tilde{F}'$ and F' do not cross after $\underline{W}$. If $\underline{W} \geq W_{nj}$, then the first step already shows that $\tilde{F}' > F'$ after $\underline{W}$. If $\underline{W} < W_{nj}$, we show below that $\tilde{F}'(W) > F'(W)$ for all $W \in (\underline{W}, W_{nj}]$, which, together with the first step, implies that $\tilde{F}'(W) > F'(W)$ for all $W > \underline{W}$.

1. $\tilde{F}'(W) > F'(W)$ for W slightly above $\underline{W}$. If $\tilde{F}''(\underline{W}) > F''(\underline{W})$, this property is obvious. If $\tilde{F}''(\underline{W}) = F''(\underline{W})$, then $\tilde{U}(\underline{W}) + \frac{\tilde{\lambda}}{r}\tilde{S}(\underline{W}) = U(\underline{W}) + \frac{\lambda}{r}S(\underline{W})$ and $a(\underline{W}) = \tilde{a}(\underline{W})$. It follows from (34), $\tilde{\lambda} > \lambda$, and $W > J(W) = \tilde{J}(W)$ that $\tilde{F}'''(\underline{W}) > F'''(\underline{W})$. This implies $\tilde{F}'(W) > F'(W)$ for W slightly above $\underline{W}$.

2. By contradiction, suppose $\tilde{F}'(W) \leq F'(W)$ for some $W \in (\underline{W}, W_{nj}]$, and define $\bar{W}$ as the second crossing point: $\bar{W} \equiv \min\{W \in (\underline{W}, W_{nj}] : \tilde{F}'(W) = F'(W)\}$. We have $\tilde{F}' > F'$ for $W \in (\underline{W}, \bar{W})$. It follows from $\tilde{F}''(\underline{W}) \geq F''(\underline{W})$ that $\tilde{U}(\underline{W}) + \frac{\tilde{\lambda}}{r}\tilde{S}(\underline{W}) \geq$

$U(\underline{W}) + \frac{\lambda}{r} S(\underline{W})$. Therefore,

$$\begin{aligned}
\tilde{U}(\bar{W}) + \frac{\tilde{\lambda}}{r}\tilde{S}(\bar{W}) &= \tilde{U}(\underline{W}) + \frac{\tilde{\lambda}}{r}\tilde{S}(\underline{W}) + (\tilde{U}(\bar{W}) - \tilde{U}(\underline{W})) + \frac{\tilde{\lambda}}{r}(\tilde{S}(\bar{W}) - \tilde{S}(\underline{W})) \\
&\geq U(\underline{W}) + \frac{\lambda}{r}S(\underline{W}) + (\tilde{U}(\bar{W}) - \tilde{U}(\underline{W})) + \frac{\tilde{\lambda}}{r}(\tilde{S}(\bar{W}) - \tilde{S}(\underline{W})) \\
&> U(\underline{W}) + \frac{\lambda}{r}S(\underline{W}) + (U(\bar{W}) - U(\underline{W})) + \frac{\tilde{\lambda}}{r}(S(\bar{W}) - S(\underline{W})) \\
&> U(\underline{W}) + \frac{\lambda}{r}S(\underline{W}) + (U(\bar{W}) - U(\underline{W})) + \frac{\lambda}{r}(S(\bar{W}) - S(\underline{W})) \\
&= U(\bar{W}) + \frac{\lambda}{r}S(\bar{W}),
\end{aligned} \tag{36}$$

where the first inequality follows from $\tilde{U}(\underline{W}) + \frac{\tilde{\lambda}}{r}\tilde{S}(\underline{W}) \geq U(\underline{W}) + \frac{\lambda}{r}S(\underline{W})$, the second from

$$\begin{aligned}
(\tilde{U}(\bar{W}) - \tilde{U}(\underline{W})) - (U(\bar{W}) - U(\underline{W})) &= (\tilde{U}(\bar{W}) - U(\bar{W})) - (\tilde{U}(\underline{W}) - U(\underline{W})) \\
&= (\tilde{F}(\bar{W}) - F(\bar{W})) - (\tilde{F}(\underline{W}) - F(\underline{W})) \\
&= \int_{\underline{W}}^{\bar{W}} (\tilde{F}'(W) - F'(W))dW > 0
\end{aligned}$$

and

$$\begin{aligned}
(\tilde{S}(\bar{W}) - \tilde{S}(\underline{W})) - (S(\bar{W}) - S(\underline{W})) &= (\tilde{S}(\bar{W}) - S(\bar{W})) - (\tilde{S}(\underline{W}) - S(\underline{W})) \\
&= (\tilde{F}(\bar{W}) - F(\bar{W})) - (\tilde{F}(\underline{W}) - F(\underline{W})) \\
&= \int_{\underline{W}}^{\bar{W}} (\tilde{F}'(W) - F'(W))dW > 0,
\end{aligned}$$

and the third inequality from $\bar{W} \leq W_{nj}$, $S(\bar{W}) - S(\underline{W}) = \int_{\underline{W}}^{\bar{W}} F''(W)(J(W) - W)dW > 0$, and $\tilde{\lambda} > \lambda$. But $\tilde{F}''(\bar{W}) \leq F''(\bar{W})$ implies $\tilde{U}(\bar{W}) + \frac{\tilde{\lambda}}{r}\tilde{S}(\bar{W}) \leq U(\bar{W}) + \frac{\lambda}{r}S(\bar{W})$, contradicting (36).

QED

A.8 Proof of Lemma 1

For the expected duration T, define $H = \int_0^\infty 1_{s<\theta}1_{s<\tau}ds$, where θ is the arrival time of the job destruction shock and $\tau = \min\{t : W_t \notin (0, W_{gp})\}$. Define a martingale H_t as $H_t = \mathbb{E}_t[H] = \int_0^t \mathbb{E}_t[1_{s<\theta}]1_{s<\tau}ds + \mathbb{E}_t[1_{t<\theta}]\mathbb{E}_t[\int_t^\infty 1_{s<\theta}1_{s<\tau}ds | t < \theta] = \int_0^t e^{-\lambda s}1_{s<\tau}ds + e^{-\lambda t}1_{t<\tau}T(W_t)$. For $t < \tau$, the drift of H_t is

$$e^{-\lambda t}\left(1 + T'(W)((r+\lambda)(W-u) + rh) + \frac{1}{2}T''(W)(r\sigma Y)^2 - \lambda T(W)\right),$$

which must be zero. For the exit probability functions P, the proof is very similar.

QED

A.9 Proof of Proposition 7

We organize the proof into three lemmas. Lemma 7 starts out by providing auxiliary results. Lemma 8 shows that $\lim_{\lambda\to\infty} P_0(W^*) > 0$. Lemma 9 shows that $\lim_{\lambda\to\infty} P_{gp}(W^*) = 0$.

Notation: Let us denote by φ the absolute value of the second derivative of F_{sep} at $W = 0$. That is, $\varphi \equiv -F''_{\text{sep}}(0)$. By assumption, we have $\varphi > 0$.

Lemma 7

1. *$F(W) \leq F_{\text{sep}}(W) + \frac{r}{r+\lambda}\bar{A}$ for all $W \in [0, W_{gp}]$.*
2. *Define $m \equiv \sqrt{\frac{4r\bar{A}}{\varphi}}$. Then, $W^* \leq \frac{m}{\sqrt{\lambda}}$ for sufficiently large λ, which implies*
$$\lim_{\lambda\to\infty} W^* = 0. \tag{37}$$
3. *$\lim_{\lambda\to\infty} F'(0) = 0$.*
4. *W_{gp} is decreasing in λ, but*
$$M \equiv \lim_{\lambda\to\infty} W_{gp} > 0. \tag{38}$$
5. *There exists $n > 0$ such that the drift of $W_t \in (0, W_{gp})$, $\mu(W_t)$, satisfies*
$$\mu(W_t) \leq n\sqrt{\lambda} \tag{39}$$
for sufficiently large λ.

Proof

1. Let (A_t, C_t, J_t) be an optimal contract starting from $W_0 = W \in [0, W_{gp}]$. We have
$$\begin{aligned}
F(W_0) &= \mathbb{E}\left[r\int_0^{\hat{\tau}} e^{-rt}(A_t - C_t)dt + e^{-r\hat{\tau}}F_{\text{sep}}(J_{\hat{\tau}})\right] \\
&\leq \frac{r}{r+\lambda}\bar{A} + \mathbb{E}\left[r\int_0^{\hat{\tau}} e^{-rt}F_0(u(C_t))dt + e^{-r\hat{\tau}}F_{\text{sep}}(J_{\hat{\tau}})\right] \\
&\leq \frac{r}{r+\lambda}\bar{A} + F_{\text{sep}}\left(\mathbb{E}\left[r\int_0^{\hat{\tau}} e^{-rt}u(C_t)dt + e^{-r\hat{\tau}}J_{\hat{\tau}}\right]\right) \\
&\leq \frac{r}{r+\lambda}\bar{A} + F_{\text{sep}}\left(\mathbb{E}\left[r\int_0^{\hat{\tau}} e^{-rt}(u(C_t) - h(A_t))dt + e^{-r\hat{\tau}}J_{\hat{\tau}}\right]\right) \\
&= \frac{r}{r+\lambda}\bar{A} + F_{\text{sep}}\left(W_0\right),
\end{aligned}$$

where the second inequality follows from $F_0 \le F_{\text{sep}}$ and the concavity of F_{sep}, the third inequality follows from $h(A_t) \ge 0$ and the monotonicity of F_{sep}, and the last equality follows from the promise-keeping constraint.

2. Pick a large $\bar{\lambda}$ so that $F_{\text{sep}}(\frac{m}{\sqrt{\lambda}}) < F_{\text{sep}}(0) - \frac{1}{4}\varphi\frac{m^2}{\lambda}$ for all $\lambda \ge \bar{\lambda}$. This is feasible because $F_{\text{sep}}(W) \approx F_{\text{sep}}(0) - \frac{1}{2}\varphi W^2$ for small W. If $\lambda \ge \bar{\lambda}$, then

$$\begin{aligned} F\left(\frac{m}{\sqrt{\lambda}}\right) &\le \frac{r}{r+\lambda}\bar{A} + F_{\text{sep}}\left(\frac{m}{\sqrt{\lambda}}\right) \\ &< \frac{r}{r+\lambda}\bar{A} + F_{\text{sep}}(0) - \frac{1}{4}\varphi\frac{m^2}{\lambda} \\ &= \frac{r}{r+\lambda}\bar{A} + F_{\text{sep}}(0) - \frac{1}{4}\varphi\frac{4r\bar{A}}{\varphi\lambda} \\ &< F_{\text{sep}}(0). \end{aligned}$$

   It follows from $F(W^*) > F_{\text{sep}}(0) > F\left(\frac{m}{\sqrt{\lambda}}\right)$ that $W^* < \frac{m}{\sqrt{\lambda}}$ for all $\lambda \ge \bar{\lambda}$.

3. We have

$$-F''(W) = \max_{a\ge 0} \frac{a + F'(W)h(a) - U(W) - \frac{\lambda}{r}S(W)}{\frac{1}{2}r\sigma^2(h'(a))^2} \le \max_{a\ge 0} \frac{a + Zh(a)}{\frac{1}{2}r\sigma^2(h'(a))^2},$$

   where Z equals $F'(0)$ under $\lambda = 0$. Hence,

$$F'(0) = F'(0) - F'(W^*) = \int_0^{W^*} -F''(W)dW \le \max_{a\ge 0} \frac{a + Zh(a)}{\frac{1}{2}r\sigma^2(h'(a))^2}W^*.$$

   It follows from (37) that $\lim_{\lambda\to\infty} F'(0) = 0$.

4. Pick a small $\underline{w}$ such that $-\max_a \frac{a+F'_{\text{sep}}(W)h(a)-U(W)}{\frac{1}{2}r\sigma^2 h'(a)^2} < F''_{\text{sep}}(W)$ for all $W \in [0, \underline{w}]$. This is feasible because $\lim_{W\to 0} F'_{\text{sep}}(W) = 0$, $\lim_{W\to 0} U(W) = F_{\text{sep}}(0)$, and we have assumed that $-\max_a \frac{a-F_{\text{sep}}(0)}{\frac{1}{2}r\sigma^2 h'(a)^2} < F''_{\text{sep}}(0)$ so a non-degenerate solution F exists.

   We show $W_{gp} > \underline{w}$ for all λ. It follows from $F(W_{gp}) = F_{\text{sep}}(W_{gp})$, $F'(W_{gp}) = F'_{\text{sep}}(W_{gp})$, and $F''(W_{gp}) \ge F''_{\text{sep}}(W_{gp})$ that

$$\begin{aligned} F''_{\text{sep}}(W_{gp}) \le F''(W_{gp}) &= \min_a \frac{-a - F'(W_{gp})h(a) + U(W_{gp}) + \frac{\lambda}{r}S(W_{gp})}{\frac{1}{2}r\sigma^2 h'(a)^2} \\ &= -\max_a \frac{a + F'_{\text{sep}}(W_{gp})h(a) - U(W_{gp})}{\frac{1}{2}r\sigma^2 h'(a)^2}, \end{aligned}$$

   which implies $-\max_a \frac{a+F'_{\text{sep}}(W_{gp})h(a)-U(W_{gp})}{\frac{1}{2}r\sigma^2 h'(a)^2} \ge F''_{\text{sep}}(W_{gp})$. Therefore, $W_{gp} > \underline{w}$.

5. By (9), the drift of W_t is $\mu(W_t) = r(W_t - u(C_t) + h(A_t)) + \lambda(W_t - J(W_t))$. We have

$$r(W_t - u(C_t) + h(A_t)) \le r(W^*_{gp} - 0 + h(\bar{A})),$$

where the right-hand side is a constant. Proposition 3 shows that $\Delta(W) = J(W) - W$ has a global minimum at $W = W^*$. We thus have

$$\lambda(W_t - J(W_t)) \le \lambda(W^* - J(W^*)) = \lambda W^* \le \lambda \frac{m}{\sqrt{\lambda}} = m\sqrt{\lambda},$$

where the last inequality follows from part 2 of Lemma 7. We can therefore find a sufficiently large n such that (39) holds.

■

Recall that $P_0(W)$ denotes the probability that the process W_t started at W reaches 0 before a job destruction shock arrives and before W_t reaches W_{gp}. The next lemma shows that $P_0(W^*)$ remains strictly positive even in the limit as $\lambda \to \infty$. In particular, it shows that $P_0(W)$ is bounded below by $\tilde{p}(W)$ defined by

$$\tilde{p}(W) \equiv p(W) - p(M), \tag{40}$$

where M is defined in (38), and

$$p(W) \equiv \exp\left(-\frac{n + \sqrt{n^2 + 2(r\sigma\gamma)^2}}{(r\sigma\gamma)^2}\sqrt{\lambda}W\right)$$

with γ defined in (1), and with n being a constant sufficiently large for (39). As shown in Lemma 7, $M > 0$. As assumed in (1), $\gamma > 0$. Clearly, $\tilde{p}$ and p are strictly decreasing and strictly convex.

Lemma 8 *$P_0(W) \ge \tilde{p}(W)$ for all $W \in [0, M]$. In particular, if $W = W^*$, then $P_0(W^*) \ge \tilde{p}(W^*) = p(W^*) - p(M) > 0$. Furthermore, $\lim_{\lambda\to\infty} P_0(W^*) \ge \lim_{\lambda\to\infty}(p(W^*) - p(M)) > 0$.*

Proof First, $p(W)$ satisfies $p(0) = 1$ and

$$\lambda p(W) = p'(W) n\sqrt{\lambda} + \frac{1}{2}p''(W)(r\sigma\gamma)^2,$$

which can be confirmed by differentiation. By (40), thus, $\tilde{p}$ satisfies $\tilde{p}(0) < 1$, $\tilde{p}(M) = 0$, and

$$\lambda\tilde{p}(W) < \tilde{p}'(W) n\sqrt{\lambda} + \frac{1}{2}\tilde{p}''(W)(r\sigma\gamma)^2. \tag{41}$$

Second, we show that, for sufficiently large λ, $e^{-\lambda t}\tilde{p}(W_t)$ is a submartingale. By Ito's lemma, the drift of $e^{-\lambda t}\tilde{p}(W_t)$ is

$$-\lambda\tilde{p}(W_t) + \tilde{p}'(W_t)\mu(W_t) + \frac{1}{2}\tilde{p}''(W_t)\nu(W_t)^2,$$

where $\mu(W_t)$ and $\nu(W_t)$ are the drift and volatility of W_t, given in, respectively, (9) and (10). By (39), for sufficiently large λ, we have $\mu(W_t) \leq n\sqrt{\lambda}$. Also, $\nu(W_t)^2 = (r\sigma h'(a(W_t)))^2 \geq (r\sigma h'(0))^2 = (r\sigma\gamma)^2$. With $\tilde{p}' < 0$, and $\tilde{p}'' > 0$, these two inequalities imply that

$$-\lambda\tilde{p}(W_t) + \tilde{p}'(W_t)n\sqrt{\lambda} + \frac{1}{2}\tilde{p}''(W_t)(r\sigma\gamma)^2$$

is a lower bound on the drift of $e^{-\lambda t}\tilde{p}(W_t)$. By (41), this lower bound is strictly positive, i.e., $e^{-\lambda t}\tilde{p}(W_t)$ is a submartingale for sufficiently large λ.

Third, let τ_M denote the first exit of W_t from $(0, M)$, a stopping time. Since, for sufficiently large λ, $e^{-\lambda t}\tilde{p}(W_t)$ is a submartingale, we have $\tilde{p}(W_0) \leq \mathbb{E}[e^{-\lambda\tau_M}\tilde{p}(W_{\tau_M})] < P_0(W_0)$.

Last, we show $\lim_{\lambda\to\infty}(p(W^*) - p(M)) > 0$. Indeed, since $W^* \leq \frac{m}{\sqrt{\lambda}}$, we have

$$\begin{aligned}
\lim_{\lambda\to\infty} p(W^*) &\geq \lim_{\lambda\to\infty} p\big(\frac{m}{\sqrt{\lambda}}\big) \\
&= \lim_{\lambda\to\infty} \exp\left(-\frac{n + \sqrt{n^2 + 2(r\sigma\gamma)^2}}{(r\sigma\gamma)^2} m\right) \\
&= \exp\left(-\frac{n + \sqrt{n^2 + 2(r\sigma\gamma)^2}}{(r\sigma\gamma)^2} m\right) \\
&> 0,
\end{aligned}$$

while

$$\lim_{\lambda\to\infty} p(M) = \lim_{\lambda\to\infty} \exp\left(-\frac{n + \sqrt{n^2 + 2(r\sigma\gamma)^2}}{(r\sigma\gamma)^2}\sqrt{\lambda}M\right) = 0.$$

■

The last lemma in this proof shows the vanishing probability of an exit at W_{gp}.

Lemma 9 $\lim_{\lambda\to\infty} P_{gp}(W^*) = 0$.

Proof Let $\tilde{W}_t$ be the solution to (8) for all $t \geq 0$, i.e., W_t is defined both before and after θ. First, we show that $F'(\tilde{W}_t)$ is a martingale. Differentiating the HJB equation (11) with respect to W yields

$$0 = F''(W)r\Big(W - u(c) + h(a) - \frac{\lambda}{r}\Delta\Big) + \frac{1}{2}F'''(W)r^2\sigma^2Y^2,$$

which means the drift of $F'(\tilde{W}_t)$ is zero.

Second, suppose $\tilde{W}_0 = W^*$. That $F'(\tilde{W}_t)$ is a martingale implies $F'(W^*) = \tilde{P}_0(W^*)F'(0) + \tilde{P}_{gp}(W^*)F'(W_{gp})$, where $\tilde{P}_0(W^*)$ and $\tilde{P}_{gp}(W^*)$ are the probabilities of $\tilde{W}_t$ reaching 0 and W_{gp}, respectively. Then

$$\tilde{P}_{gp}(W^*) = \frac{F'(W^*) - \tilde{P}_0(W^*)F'(0)}{F'(W_{gp})} = \frac{0 - \tilde{P}_0(W^*)F'(0)}{F'_{\text{sep}}(W_{gp})}.$$

As $\lambda \to \infty$, the limit of the above is 0 because $\lim_{\lambda\to\infty} F'(0) = 0$ by Lemma 7.3, and $\lim_{\lambda\to\infty} F'_{\text{sep}}(W_{gp}) = F'_{\text{sep}}(M) \neq 0$.

Finally, $\lim_{\lambda\to\infty} P_{gp}(W^*) = 0$ because $P_{gp} < \tilde{P}_{gp}$. ■

This concludes the proof of Proposition 7.

QED

References

Almazan, A. and J. Suarez (2003). Entrenchment and severance pay in optimal governance structures. *The Journal of Finance 58*(2), 519–547.

Anderson, R. W., M. C. Bustamante, S. Guibaud, and M. Zervos (2018). Agency, firm growth, and managerial turnover. *Journal of Finance 73*(1), 419–464.

Boeri, T., P. Garibaldi, and E. R. Moen (2017). Inside severance pay. *Journal of Public Economics 145*, 211–225.

Buchinsky, M., D. Fougere, F. Kramarz, and R. Tchernis (2010). Interfirm mobility, wages and the returns to seniority and experience in the United States. *Review of Economic Studies 77*(3), 972–1001.

Comin, D. and S. Mulani (2006). Diverging trends in aggregate and firm volatility. *Review of Economics and Statistics 88*(2), 374–383.

Cozzi, M. and G. Fella (2016). Job displacement risk and severance pay. *Journal of Monetary Economics 84*, 166–181.

DeMarzo, P. M., D. Livdan, and A. Tchistyi (2014). Risking other people's money: Gambling, limited liability, and optimal incentives. *Stanford Working Paper No.3149*.

Edmans, A. and X. Gabaix (2016). Executive compensation: A modern primer. *Journal of Economic Literature 54*(4), 1232–87.

Epiq (2022). Pandemic bankruptcy battles: Looking back and beyond. `https://www.jdsupra.com/legalnews/pandemic-bankruptcy-battles-looking-1066272/`.

He, Z. (2012). Dynamic compensation contracts with private savings. *Review of Financial Studies 25*(5), 1494–1549.

Hoffmann, F. and S. Pfeil (2010). Reward for luck in a dynamic agency model. *Review of Financial Studies 23*(9), 3329–3345.

Inderst, R. and H. M. Mueller (2010). Ceo replacement under private information. *The Review of Financial Studies 23*(8), 2935–2969.

Lee Hecht Harrison (2018). Severance and separation benefits benchmark study.

Li, R. (2017). Dynamic agency with persistent observable shocks. *Journal of Mathematical Economics 71*, 74–91.

Mortensen, D. T. and C. A. Pissarides (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies 61*(3), 397–415.

Piskorski, T. and A. Tchistyi (2010). Optimal mortgage design. *Review of Financial Studies 23*(8), 3098–3140.

Pissarides, C. A. (1985). Short-run equilibrium dynamics of unemployment, vacancies, and real wages. *American Economic Review 75*(4), 676–690.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economic Studies 75*(3), 957–984.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review 95*(1), 25–49.

Spencer, J. (2021). Amid pandemic, some executives still exit with eye-popping severance packages. `https://jacksonspencerlaw.com/top-pandemic-severance-packages/`.

Topel, R. (1991). Specific capital, mobility, and wages: Wages rise with job seniority. *Journal of Political Economy 99*(1), 145–176.

Wang, C. (2011). Termination of dynamic contracts in an equilibrium labor market model. *Journal of Economic Theory 146*(1), 74–110.

Wong, T.-Y. (2019). Dynamic agency and endogenous risk-taking. *Management Science 65*(9), 4032–4048.